Title: Methodological approaches and related challenges associated with the determination of critical power and *W’*

Running title: Critical power and *W’* determination

Word count: 6500 words (not including abstract, tables and figures, and reference list).

Abstract

The relationship between exercise intensity and time to task-failure (*P-T* relationship) is hyperbolic, and characterised by its asymptote (critical power, CP) and curvature constant (*W'*). The determination of these parameters is of interest for researchers and practitioners, but the testing protocol for CP and *W'* determination has not yet been standardised. Conventionally, a series of constant work-rate tests (CWR) to task-failure have been used to construct the *P-T* relationship. However, the duration, number, and recovery between predictive CWR, and the mathematical model (hyperbolic or derived linear models) are known to affect CP and *W'*. Moreover, repeating CWR may be deemed as a cumbersome and impractical protocol. Recently, CP and *W'* have been determined in field and laboratory settings using time-trials, but the validity of these methods has raised concerns. Alternatively, a 3-min all-out test (3MT) has been suggested, as it provides a simpler method for the determination of CP and *W'*, whereby power output at the end of the test represents CP, and the amount of work performed above this end-test power equates to *W'.* However, the 3MT still requires an initial incremental test, and may overestimate CP. The aim of this review is, therefore, to appraise current methods to estimate CP and *W'*, providing guidelines and suggestions for future research where appropriate.

Key words: Exercise tolerance; Exercise domains; Fitness testing; Performance; Fatigue

1. Introduction

The relationship between exercise intensity and time to task-failure (T_{lim}) (i.e. the *P-T* relationship) has received extensive research attention. The first attempts to model the *P-T* relationship date back to the beginning of the 20th century when Kennelly (69) and Hill (50) studied the speed of humans and animals over various distances. However, Scherrer and Monod (95) formally described the *P-T* relationship as hyperbolic in a single-joint muscle action. The *P-T* relationship appears to be highly conserved, and has subsequently been observed in various forms of whole body exercise, in individuals with different levels of fitness, and across animal species (90).

The hyperbolic *P-T* relationship is characterised by two parameters. The asymptote of the hyperbola is defined as critical power (CP), and the curvature constant is notionally abbreviated as *W'*. Briefly, it has been suggested that CP demarcates the highest exercise intensity at which metabolic and systemic responses attain a steady state (61,90,91). Where power is directly measurable (e.g. cycling), CP is typically expressed as a mechanical power output (PO). However, factors which affect the relationship between oxygen consumption ($\dot{V}O_2$) and PO, such as cadence, are known to also affect CP (8), and indeed some authors have proposed to use the term 'critical intensity' and to express CP as a $\dot{V}O_2$ equivalent (118). However, as expressing CP as a PO may be more applicable (86) and freely chosen cadence is relatively consistent within individuals (47), this review will consider CP as a mechanical PO. With regards to *W'*, it represents the amount of work that can be performed above CP, and was originally considered to represent anaerobic energy production (51,81). However, it is now accepted that the precise aetiology of *W'* is more complex, and affected by factors such as accumulation/depletion of intramuscular substrates and fatigue-related metabolites (90). Further details on the aetiology of CP and *W'* are discussed elsewhere (59,90,108).

The determination of CP and *W'* is of interest to researchers and practitioners alike. For instance, prescribing exercise intensities relative to CP may elicit a more homogenous response than other approaches to normalise the intensity of exercise, such as a percentage of maximum oxygen consumption ($\dot{V}O_{2max}$) (4,71,74). Secondly, exercise within the 'severe' domain, above CP, results in a progressive depletion of *W'*, so that when *W'* is depleted, exercise is either terminated or the intensity reduced to <CP. The determination of CP and *W'* therefore allows prediction of the time to reach T_{lim} during exercise above CP. These predictions are typically within 15% of the actual T_{lim} , and actual and predicted T_{lim} are strongly correlated ($r \geq 0.87$) (29,41,62,68,84,87,114). Thirdly, CP is strongly associated with endurance performance, and it has been shown to account for 69-86% of the variance in sporting events lasting ~2.2 to ~59 min (17,20,70,99). Similarly, running events lasting longer than 1 h, such as the marathon, are also strongly correlated with the running equivalents of CP (termed critical speed (CS)), and completed at an intensity close to, but fractionally below, CP (41,59).. Moreover, the combination of CS and the running equivalent of *W'* (*D'*) predicts 5000-m running performance within 1% (85). Finally, with the advantages of the aforementioned applications, it is not surprising that the *P-T* relationship has been used to evaluate and monitor performance, and proposed as a tool for anti-doping (37,93,116).

The determination of CP and *W'*, however, is not standardised. In most laboratories, CP and *W'* have been determined using a series of square-wave constant work-rate tests to task-failure (CWR), in which T_{lim} is recorded. These CWR are usually interspersed with 24 h of recovery, making this method cumbersome and impractical. Several attempts have been made to simplify the protocol, including reducing the number of CWR required, or shortening the 24-h recovery duration between CWR. In addition, advancements in the development of power meters and ergometers have facilitated the determination of CP and *W'* using time-trials (TT), both in the field and the laboratory. Alternatively, CP and *W'* may be determined using a 3-min all-out test (3MT), whereby the mean PO during the final 30 s of the test represents CP, and the amount of work performed above that mean end-test PO represents *W'*. However, the above approaches have limitations, and there are methodological challenges that need to be considered. The estimation of CP and *W'* is influenced by the testing protocol and, as a result, research findings between studies are difficult to compare.

This review aims to draw attention to these issues and, where appropriate, to state relevant recommendations for the determination of CP and *W'*.

2. Conventional approach to determine CP and *W'*: mathematical models, and duration, number, and recovery between tests.

The conventional approach to determine CP and *W'* in a laboratory setting requires the performance of 3–5 CWR, where PO and T_{lim} are recorded. From these data, total work performed (i.e. $Work = PO \times T_{lim}$) and the inverse of T_{lim} (i.e. ${T_{lim}}^{-1}$) can be calculated (Table 1); with subsequent linear and non-linear models applied to estimate CP and *W'* (43,49,51,60,81).

Figure 1 near here

Table 1 near here

PO and T_{lim} derived from each CWR can be fitted using a hyperbolic function (Figure 1A). The asymptote of the hyperbola represents CP, and the curvature constant denotes *W'*. For any given PO above CP, the duration of exercise to task-failure (i.e. T_{lim}) is determined as:

$$T_{lim} = \frac{W'}{PO - CP} \qquad [1]$$

The non-linear equation [1] can be rearranged to a linear function by plotting PO against the inverse time (${T_{lim}}^{-1}$). Here, the slope of the line represents *W'*, and the *y*-intercept represents CP (Panel 1B):

$$PO = CP + W' \times {T_{lim}}^{-1} \qquad [2]$$

An alternative linear function of the *P-T* relationship may be obtained by plotting the work accomplished in each CWR against T_{lim} (Figure 1C). The y-intercept of this line represents *W'*, and the slope represents CP:

$$Work = W' + CP \times T_{lim} \qquad [3]$$

Fitting the *P-T* relationship with a 2-parameter function (non-linear or derived linear functions) has some limitations. For instance, as T_{lim} approaches zero, PO becomes infinite. To overcome this limitation, a third parameter, k, has been introduced (80):

$$T_{lim} = \left(\frac{W'}{PO-CP}\right) + k \qquad [4]$$

where k is interpreted as the maximum instantaneous PO (PO_{max}). Hence, with the inclusion of k, as T_{lim} approaches zero, PO approaches PO_{max}. CP and *W'* can be determined from a 3-paramter model, in which k is substituted as:

$$T_{lim} = \left(\frac{W'}{PO-CP}\right) + \left(\frac{W'}{CP- PO_{max}}\right) \qquad [5]$$

Another limitation of 2-parameter models is the assumption that, for any intensity below CP, there is no contribution of *W'* at the onset of exercise. However, with a demonstrated link between CP and $\dot{V}O_2$ on-kinetics (46,83), some authors have suggested that *W'* contribution at the onset of exercise may be somewhat underestimated (60,82). Wilkie (117) proposed accounting for $\dot{V}O_2$ on-kinetics through the use of a rather fast time constant of 10 s for all individuals. While the inclusion of the time constant of $\dot{V}O_2$ on-kinetics appears to be physiologically sound, it seems a cumbersome addition and is currently not used. Further research may investigate whether the inclusion of an individually-derived time constant improves the precision of CP and *W'* estimations.

An area of concern is the test-retest reliability of the estimates of CP and *W'* derived from CWR. Using the linear T_{lim}^{-1} model (Equation [2]), the coefficient of variation (CV) and correlation coefficient (r) of CP have been reported at 3% and 0.96, respectively; whereas the corresponding values for *W'* were 10.3% and 0.79, respectively (44). It is worth noting that a 10-15% variability in T_{lim} has been observed in CWR (5,72,82). A large variation in *W'* may occur as a result of the nature of the mathematical model, since small changes in T_{lim} during exhaustive CWR have a negligible effect on CP, but a much larger effect on *W'* (93,105,107). Nonetheless, the test-retest reliability seems to be poorer for *W'* than CP using other methodological approaches (e.g. TT or all-out tests, see discussion below). Furthermore, studies comparing different approaches to determine CP and *W'* typically report a closer agreement between methods for estimating CP than for *W'* (e.g. (65,85,96,103,109,119)), although a high reliability for both parameter estimates (ICC of 0.94 and 0.95 for CP and *W'*, respectively) was reported after a familiarization trial when using TT under controlled laboratory conditions (103). Overall, however, *W'* appears to exhibit a greater variability than CP, though the reason(s) for this phenomenon are not yet completely understood.

2.1. Effect of the mathematical modelling on CP and *W'* estimations

The equations described above typically fit the data with a high degree of accuracy ($R^2 \geq 0.82$) (14,23,43). However, they result in different estimations of CP and *W'*, even though some of these equations [1-3] are mathematically equivalent (14,19,20,22,23,43,56,94). Depending on the model, estimations of CP typically are, from highest to lowest, in the following order: linear T_{lim}^{-1} model (equation [2]), linear total work model (equation [3]), 2-parameter hyperbolic model (equation [1]), and 3-parameter model (equation [5]); with estimations of *W'* following the reverse order (Figure 2). It is important to note that in some studies no differences between mathematical models were reported (e.g. (19,31,105)). Nonetheless, irrespective of whether estimations of CP derived from different mathematical models reach statistical significance, large T_{lim} differences have been observed during exercise at respective CP intensities, ranging ~20-60 min (21,23,51,77,85,87).

The question of which mathematical model should be used to determine CP and *W'* remains unresolved. The 3-parameter model consistently produces lower estimates of CP and greater estimates of *W'* than 2-parameter models (14,20,22,28,43). Furthermore, the 3-parameter protocol, suggested by Morton (80), requires a relatively large number of trials, including some with low (<1 min) and high (>15 min) T_{lim}, which in turn can affect the estimation of CP and *W'* (see section 2.2). Moreover, the 3-parameter model may produce non-physiological estimates of PO_{max}, and the parameter exhibits large inter-subject variability (28,43,80). These issues may explain why most recent studies have indeed used 2-parameter models (e.g.

(61,63,79,91)). An alternative approach has been proposed by Hill (51), and recently adopted by some researchers (18,19,101), whereby the model producing the lowest standard error of estimate (SEE) is used. We therefore recommend that the *P-T* relationship should be characterised with the 2-parameter model that results in the lowest SEE.

2.2. Effect of duration of predictive trials on CP and *W'*

The characteristics of the tests used to define the *P-T* relationship have a profound effect on CP and *W'* estimates. For instance, the duration of CWR is known to affect CP and *W'* (16,26,57,75,102,106,115). If data from five tests to task-failure is rearranged, and only the three tests with the shortest durations are considered, CP has been shown to be 14-20% greater than that derived from the three longest durations, irrespective of the overall range of duration of all five exhaustive CWR (16,57). Moreover, *W'* appears to be notably more sensitive to the duration of the trials, with the three shortest exhaustive trials producing *W'* estimates ~70% greater than those derived from the three longest trials (16). The effect of trial duration on CP and *W'* is shown in Figure 3.

Scherrer and Monod (95) stipulated that the work-T_{lim} relationship (equation [3]) loses linearity for exercise durations <2 min, with di Prampero (92) specifying that the range of test durations should be such that $\dot{V}O_{2max}$ is elicited, and that *W'* is fully depleted during each trial. However, the first requirement is not always verified (48,53,75,81), and a complete depletion of *W'* may be difficult to assess. At very high intensities (i.e. short T_{lim}), *W'* may contribute more than the model predicts due to the relatively slow increase in $\dot{V}O_2$ (16,81,107). Moreover, at such high intensities, it is possible that exercise terminates before $\dot{V}O_{2max}$ has been reached (27,52,92,105). Therefore, trials with a T_{lim} <2 min should be considered too short and not included in the determination of CP and *W'* (16,60,91,92). On the other hand, exercise performed above CP and continued for >2 min should lead to maximal values of $\dot{V}O_2$ and blood lactate concentration (19,25,88). However, some studies have reported that $\dot{V}O_2$ did not reach its maximum at task-failure during the longest predictive trials, which corresponded to intensities slightly (~10%) above CP (11,94). The reason(s) for this phenomenon remain unknown, but it is likely to be multifactorial, including physiological and/or psychophysiological factors (1,11,94). Therefore, it is recommended that exhaustive trials which result in T_{lim} >15 min should be avoided as $\dot{V}O_{2max}$ may not be reached. Furthermore, whenever possible, and at least for research purposes, we recommend that the attainment of $\dot{V}O_{2max}$ should be verified for all predictive trials.

The range in the duration of the trials should also be considered when investigating alternative testing protocols (i.e. duration of criterion versus experimental trials) (104). In order to minimise such effects, it is now common that CP and *W'* are determined from trials with T_{lim} ranging between 2 and 15 min, with a minimum of at least 5 min between the longest and shortest trial (e.g. (67,105,112)). Nonetheless, it has been shown recently that the duration of the predictive trials may still affect the estimation of CP and *W'*, even when these trials are performed within the recommended T_{lim} range of 2-15 min. Triska et al. (102) determined CS and *D'* from two protocols: three TT of 12, 7, and 3 min and three TT of 10, 5, and 2 min. The former protocol resulted in ~3% lower CS and ~14% higher *D'* compared to the latter protocol. It is unclear if these findings can be extrapolated to other forms of exercise such as cycling, but these data suggest that a consistent protocol should be used to assess or monitor performance using the CP model.

In summary, 2-15 min is the recommended duration of trials, and exhaustive trials resulting in a T_{lim} <2 min or >15 min should be excluded from calculations. The specific duration of predictive trials should also be considered, even if the overall range of durations falls within the target of 2-15 min. Alternatively, research investigating the effects of a treatment may employ the same duration (i.e. TT). Furthermore, the attainment of $\dot{V}O_{2max}$ should be verified wherever possible before including respective trials in the calculation of CP and *W'*.

2.3. Effect of the number of trials on CP and *W'*

Critical power and *W'* can be determined from just two trials. Indeed, CP determined from two exhaustive trials with relatively different T_{lim} (>15 min) was only ~1.1% greater than that determined using four trials (55). More recently, Simpson and Kordi (97) determined CP and *W'* in experienced cyclists using a protocol consisting of two laboratory-based TT of 3 and 12 min, interspersed with 40 min of passive rest. The authors noted that, after two familiarisation sessions, the addition of a third trial of intermediate duration (5 min) did not affect CP or *W'*. A potential limitation of this approach is that using only two exhaustive trials always results in a perfect fitting of the model, and therefore SEE cannot be determined. Instead, to ensure a high quality of the model, particularly for research purposes, the *P-T* relationship is most commonly determined from three or more CWR to task-failure (51). Indeed, a recent approach proposes performing trials until the model falls within a certain SEE; for example, less than 2% (36,40,102) or 5% (18,19) for CP, and less than 10% for *W'* (18,19,36,40,102). In summary, using only two exhaustive trials may seem an attractive option to determine CP and *W'* in the interest of a short protocol. However, where possible and at least for research purposes, we recommend using three or more trials, so that the *P-T* relationship provides estimates within predetermined SEE's for CP and *W'*.

2.4. Duration of the recovery between exhaustive trials

The duration of the recovery between exhaustive trials is usually at least 24 h, which makes the determination of the *P-T* relationship cumbersome. To address this issue, some authors have investigated whether a shorter recovery between trials affects CP and/or *W'* (15,45,63,85,97,105). Karsten et al. (64) compared the conventional 24 h method with two experimental recovery durations of 3 h and 30 min. The authors observed that, in comparison with the standard 24-h-recovery protocol, the two shorter recovery protocols were sufficient to not affect CP (prediction error of 2.5% and 3.7% for the 3 h and 30 min recovery protocols, respectively, compared to 24 h). However, the prediction error inherent in the experimental protocols was higher for *W'* (25.6% and 32.9% for the 3-h and the 30-min protocols, respectively). The authors proposed a couple of reasons to explain these findings. Firstly, the shorter recovery protocols might have led to only a partial reconstitution of *W'*; although *W'* may be restored within ~25 min following exhaustive exercise (33,39,98). Secondly, high-intensity exercise can affect the $\dot{V}O_2$ on-kinetics and increase (i.e. 'prime') performance in subsequent exercise performed up to 45 min after the initial bout (3,24). However, Karsten et al. (63) more recently showed that $\dot{V}O_2$ on-kinetics were not significantly different between repeated CWR and TT following a 60-min recovery period, suggesting that, at least for the 3-h recovery intervention, the argument does not hold. In summary, a single-day determination of CP can be achieved by reducing the inter-trial recovery time to 30 minutes. However, at present, a more conservative recovery of 60-min is preferred to determine both CP and *W'*, in order to minimise any potential priming effect and to allow for a full reconstitution of *W'*.

3. Determination of CP and *W'* using time trials under laboratory and field conditions

3.1. Laboratory and field determination of critical power and *W'*

With the popularisation of power meters PO data is readily available, which allows analysis of the *P-T* relationship in the field. For instance, PO data from elite cyclists over a competitive season have been reported for exercise durations ranging from 1 s to 4 h and, unsurprisingly, mean PO decreases nonlinearly as the duration increases (89). Indeed, a translation of laboratory-based determination of CP and *W'* into the field was attempted by Karsten et al. (65). The study compared CP and *W'* results, using three laboratory CWR (resulting in task-failure times of ~12, 4, and 2.5 min) with those determined from three track-based TT where participants had to produce the highest possible PO for 12, 7 and 3 min. All tests were performed on separate days and the authors reported a close agreement between laboratory and field CP values (prediction error of 7 W). However, field values of *W'* were ~5 kJ higher than those obtained in the laboratory, irrespective of the mathematical model used. In a follow up study (67), a shortened testing protocol (i.e. a 30 min intra-trial recovery period; see Section 2.4) was used to investigate whether CP and *W'* could be reliably determined from road PO data. The study comprised three experimental protocols and a criterion protocol to determine CP and *W'.* The criterion protocol consisted of three laboratory-based CWR interspersed with 30-min recovery; and the experimental protocols were: i) a TT field-based protocol consisting of three maximal exhaustive efforts over 12, 7 and 3 min, interspersed with 30-min recovery; ii) a field-based protocol consisting of three TT over the same durations, but interspersed with 24-h recovery; and iii) non-intentional TT maximal efforts (i.e. highest PO over the three durations obtained at any point during a single training session). The results demonstrated a high agreement for all experimental CP values with a mean prediction error of ~11, 17 and 14 W for protocols i, ii, and iii, respectively. However, results for *W'* showed an unacceptably high prediction error of ~3, 4, and 3 kJ, respectively. All experimental protocols were repeated three times with a mean within-protocol CV for CP of 2.4%, 6.5%, and 3.5%, respectively. Of note is that protocol ii is at the upper end of what is considered as acceptable reliability for physiological variables in sports science research (2,54). With regards to *W'*, only protocol iii, the non-intentional efforts, provided a relatively low CV for *W'* (~17%) when compared to protocol i (~46%) and protocol ii (~45%). Triska et al. (105) compared a single-day field test to estimate CP and *W'* (three TT of 12, 6, and 2 min) with a laboratory-based protocol using a cadence dependent (i.e. linear) mode to mimic 'real-world' exercise. The authors reported similar mean values between conditions for CP (laboratory: ~280 W vs. field: ~281 W), and a 95% LoA of -55 – 50 W. In contrast, *W'* was significantly higher under laboratory conditions (~21.6 vs. ~16.3 kJ) with a correspondingly poor agreement (95% LoA: -3.5 – 16.4 kJ) between protocols. Altogether, these data suggest that CP can be determined with reasonable precision in the field, or by simulating field conditions (i.e. using TT). However, *W'* appears to be under- (single-day approach, (105)) or over-estimated (multi-day approach, (65)) using these tests; though reasons have not yet been elucidated.

3.2. Time-trial versus constant work-rate tests

There are a number of methodological differences between laboratory- and field-based tests that need to be considered within the context of CP and *W'* determination. First, laboratory-based protocols typically use open-end tests (i.e. CWR), whereas field tests typically employ maximal effort over a fixed time or distance (i.e. TT). Time-trials exhibit less test-retest variation than CWR (72), and therefore resulting in significantly lower SEE for CP and *W'* estimates (63). Secondly, TT are self-paced, and pacing has been shown to affect the *P-T* relationship (18,62). Black et al. (18) compared estimations of CP and *W'* derived from 4-6 CWR prediction trials performed on different days with work-matched TT in the laboratory. Despite being equalled for work, mean PO was higher, and therefore T_{lim} shorter during TT, possibly due to the fast-start commonly adopted in TT (18). As a result, CP was ~7% higher using TT, whereas *W'* was not affected by the type of exhaustive trials; though there was a negative correlation (r = -0.74) between the relative change in CP and *W'* in CWR and TT (18). In contrast, Karsten et al. (63) compared non time-matched CWR with TT in the laboratory, with a recovery time of 60 min between efforts to avoid a possible $\dot{V}O_2$ priming effect evident with shorter recovery periods (see Section 2.4). The results demonstrated a low prediction error for CP (2.7%; 8 W), but a high prediction error for *W'* (18.8%; 2.5 kJ); though it is likely that the latter was influenced by the relatively short recovery period between efforts. It is also worth noting that Black et al. (18) utilised self-paced TT, where the ergometer was set in linear mode with a fixed resistance (i.e. cadence-dependent mode) allowing PO to be regulated by cadence only, whereas Karsten et al. (63) utilised self-paced TT, where the ergometer allowed PO to be self-regulated using changes in gear ratio (virtual) and cadence, in an attempt to better replicate real-world cycling. Thirdly, TT are not constrained by cadence, whereas CWR are commonly performed at a predetermined cadence (105), and pedalling rate is known to affect CP and *W'* (8,34,73,110). Fourthly, the duration of CWR is variable, whereas it can be standardised for TT. As a result, there might be differences in the duration of exhaustive trials (18), which, as discussed above, can affect CP and *W'*. Further evidence for the effects of time differences also comes from other exercise modes. In running, Galbraith et al. (45) reported that estimations of CS derived from three TT interspersed with either 30 or 60 min of passive rest between trials were not significantly different from three CWR performed in the laboratory using a multi-day protocol (typical error 0.14 $m \cdot s^{-1}$ and 0.16 $m \cdot s^{-1}$ for 30 or 60-min rest, respectively). In contrast, field-based estimations of *D'* were significantly lower (typical error 88 m and 84 m for 30 or 60-min rest protocols, respectively) than those derived from a laboratory-based test. The field-based approach also exhibited comparable test-retest variability to that obtained from the conventional laboratory-based approach (0.4% and 13% for CS and *D'*, respectively). Triska et al. (104) attempted to address the issues surrounding the values of *D'* by time-matching the laboratory and the field trial durations. The authors reported no differences and positive correlations for CS and *D'* between the two conditions, and LoA of ±0.24 $m \cdot s^{-1}$ and ±75.5 m. These studies seem to indicate that reasons other than that of trial duration are responsible for the conundrum surrounding *D'*. Fifthly, there appear to be a number of factors during field-based TT protocols that might affect CP and *W'* such as standing vs. rolling starts, overcoming inertia and acceleration, increased air resistance, or differences in terrain (78,88,105). The precise role of each of these factors warrants further investigation. On the other hand, field based-based tests can offer a more ecologically valid approach to estimate CP and *W'*. This is particularly true if CP and *W'* are to be used in the field, where the above issues of acceleration, pacing or air resistance, remain present. A final point to consider is the test-retest reliability of estimations of CP and *W'* using TT. Recently, Triska et al. (103) performed three identical TT to determine CP and *W'* using a single-day protocol with the first TT used as familiarisation. The authors noted that the CV of CP and *W'* between the familiarisation and the first subsequent TT were 4.1% and 25.3%, respectively. However, the analysis of the two consecutive TT performed after familiarisation produced closer estimates in both CP and *W'* (2.6% and 8.2%, respectively). Therefore, the authors concluded, familiarisation is advisable to determine CP and *W'* from TT using a single-day protocol.

In summary, although laboratory-based TT can be used to determine CP and *W'*, some discrepancies in the estimation of CP and, in particular, *W'* are evident. Nonetheless, and even though there are methodological differences between CWR and TT protocols, TT may be preferable over CWR, particularly if the data are to be used under field conditions. If CP and *W'* are determined from TT, performing a familiarisation trial is advisable to increase the reliability of the estimates.

4. The 3-min all-out test

The conventional approach to determine CP and *W'* requires the performance of repeated maximal efforts, which may compromise the practical application of the model. It has been hypothesised that the parameters of the *P-T* relationship may be obtained from a single all-out test. The rationale is that, at the start of all-out efforts, *W'* is heavily utilised; however, as the exercise continues and PO decreases, so does *W'.* If the duration of exercise is sufficiently long, *W'* becomes fully depleted and, therefore, the PO at or towards the end of an all-out effort should represent CP. Dekerle et al. (35) first explored this idea using an all-out effort lasting 90 s; but the authors noted that at the end of the test, PO was greater than CP, and that *W'* was not fully depleted. Burnley et al. (25) extended the duration to 180 s, and observed that the decrease in PO had stabilised in the final 30 s of the test (defined as 'end-test power output' [EP]) (Figure 4). In a follow-up study, a close agreement was reported between the conventionally determined CP and the EP obtained during a 3MT (r = 0.99; SSE = 6.4 W) (109). Moreover, the work performed above EP (WEP) was similar to *W'* (r = 0.84; SEE = 2.6 kJ). For the purpose of this review we will use CP and *W'* when referring to results derived from the conventional protocol using CWR or TT, and EP and WEP when referring to the 3MT.

The original 3MT still requires two testing days, as a prior exhaustive incremental maximal test is a prerequisite for the subsequent ergometer setting, using values of gas exchange threshold (GET), preferred cadence, and $\dot{V}O_{2max}$ (25,109). The 3MT starts with a period of unloaded cycling after which participants are instructed to accelerate their cadence up to 110–120 rpm at which point the cycle-ergometer switches into the linear mode. The linear factor is set so that at the participant's preferred cadence, the PO corresponds to halfway between GET and $\dot{V}O_{2max}$ (50%Δ; Equation [6]), which is suggested to approximate CP (25):

$$Linear\ factor = \frac{PO\ at\ 50\%\Delta}{Cadence^2} \qquad [6]$$

As fatigue develops during all-out exercise, cadence drops resulting in a decline in PO and the typical curvilinear 3MT power profile. To prevent pacing, participants are blinded to elapsed time, and strong verbal encouragement is required throughout the test. To provide reliable results, a familiarisation 3MT trial is also commonly performed, increasing the overall time required to determine EP and WEP. Performing a GXT, a familiarisation trial and the actual 3MT necessitates more than one laboratory visit, which in turns lengthens a protocol that benefits from an otherwise short testing methodology.

There are no formal criteria to verify the validity of the 3MT. However, some authors reported that PO plateaus towards the end of the 3MT, as determined using consecutive 30-s bins (25,42). It has been also reported that PO peaks within the first 10 s (109), and subsequently decreases rapidly so that >90% of WEP is depleted within the first 90 s of the test (110). In addition, as an all-out effort is required, a decrease in PO greater than 5% of EP (see discussion below on reliability) for 5 s may denote pacing and cause some reconstitution of WEP, and therefore an overestimation of this parameter. An accurate selection of the linear factor is crucial, since relatively small alterations in preferred cadence by ±10 rpm can significantly affect EP and/or WEP and end test cadence (110). To reflect the maximal (i.e. all-out) nature of the test, $\dot{V}O_2$ has been suggested to attain its maximum during a 3MT (25,42,109); and blood lactate concentration reaches >8 $mmol{\cdot}L^{-1}$ (25,110,113). In summary, the following criteria may be proposed to ensure a true 3MT all-out effort: i) a plateau in PO in the last 30 s of the test; ii) the attainment of peak PO within the first 10 s of the test; iii) rapid initial decrease of PO, so that >90% of WEP is depleted within the first 90 s of the test; iv) no decrease in PO >5% EP for >5 s during the test; v) an end-test cadence within 10 rpm of preferred cadence; vi) the attainment of $\dot{V}O_{2max}$; and vii) a blood lactate concentration >8 $mmol{\cdot}L^{-1}$. With regards to the reliability of EP and WEP, both parameters show a similar degree of reliability to those derived from the conventional testing approach. Specifically, the reliability of EP has consistently been shown to be better (CV of 3-7%) than that of the WEP (8-21%) (25,38,58,73).

4.1. Single-day alternatives of the original 3MT

As the original 3MT requires two laboratory visits, several authors have attempted to shorten or to simplify the original 3MT. For instance, Johnson et al. (58) proposed that the resistance of the 3MT may be determined relative to body mass, somewhat similar to the Wingate anaerobic test. Bergstrom et al. (10) reported that a modified 3MT, performed on a mechanically-braked ergometer, with resistances set at 4.5% body mass, could be used to determine EP and WEP. However, if the resistance was set at 3.5% body mass the modified 3MT produced different estimates of EP and WEP than those derived from the original 3MT and from the conventional approach (10); although the error was not reported, and agreement between methods was identified using a test of difference. In a similar study, Clark et al. (31) performed a 3MT on a mechanically braked ergometer using loads of 3, 4, or 5% of body mass for recreationally active, anaerobic and aerobic athletes, and endurance athletes, respectively. There were no significant differences in either EP or WEP determined from the 3MT, irrespective of whether values were determined using linear factors based upon body mass or using the conventional linear factor of 50%Δ. The authors, however, reported a large individual variation between the methods in estimates of EP and, particularly, WEP (4.2% and 39.4%, respectively). Dicks et al. (38) calculated the linear factor based on age, gender, body mass and self-reported physical activity levels. The authors reported no differences in either EP or WEP between the original 3MT and the alternative 3MT. Moreover, there were no differences between the parameters of the *P-T* relationship derived from the alternative 3MT, and those derived from three CWR using linear models (Eqs. [2,3]). However, the CV between methods was again much higher for WEP (≥ 21.8%) than for EP (≤ 4.8%) (38). In addition, Dicks et al. (38) used CWR lasting ~3, 4, and 5 min to model the *P-T* relationship; possibly overestimating CP and underestimating *W’* (see Section 2.2). Constantini et al. (33) evaluated the effects of performing the incremental test and 3MT in a single testing session. The authors reported that a 3MT performed 20 min after the incremental test resulted in EP and WEP values similar to those obtained when the 3MR and incremental test were performed over different days (SEE 5 W and 1.81 kJ for EP and WEP, respectively). Clark et al. (30) evaluated the merits of performing a 3MT on the CompuTrainer, a training ergometer often used by cyclists. The results showed a good agreement between conventional (linear work and T_{lim}^{-1} models) and 3MT approaches for determining CP and EP (2.8% and 3.1%, respectively). However, a poor agreement between WEP and *W'* derived from the linear Work-T_{lim} (CV of 24.4%) and PO-T_{lim}^{-1} (CV of 26.3%) models was also reported.

In summary, various alternatives have been proposed to simplify the conventional 3MT. Overall, alternative approaches of the 3MT discussed above seem to produce similar EP values compared to the original 3MT. However, since WEP seems to exhibit large variation, alternative protocols to the 3MT warrant caution, and as such, the conventional approach is preferred.

Most of research focusing on the 3MT has been performed in healthy and athletic populations; most likely because of the challenging nature of sustaining an all-out effort for three minutes. It is nonetheless worth noting that the 3MT has been performed by adolescents (14-15 years), who might have a reduced anaerobic fitness compared to adults (7). No significant differences were observed between the conventional and 3MT approaches to estimate CP/EP and *W'*/WEP values in adolescents; though a large variation (~20%) within-individuals prevented the 3MT and conventional approaches from being used interchangeably (6). Future research should consider whether the 3MT is a feasible option for non-athletic populations, particularly those with limited fitness.

4.2. Critical appraisal of the 3-min all-out test

Other approaches have been adopted to determine CP and *W'* using a 3MT, which provide further insight into the validity of EP and WEP for estimation of CP and *W'.* For instance, several studies have investigated the 3MT using isokinetic cycling exercise. Dekerle et al. (34) reported that the isokinetic 3MT produced measures of CP and *W'* that were not significantly different from those derived using the traditional approach; although the large intra-subject variability, in particular for WEP, led the authors to caution against the use of the isokinetic

3MT. Karsten et al. (66) reported a greater EP (~7%) and smaller WEP (~25%) derived from an isokinetic 3MT than those obtained from the conventional approach, with poor levels of agreement between these two approaches. In contrast to the above, Wright et al. (119) conducted the only study to date comparing the conventional CWR with the 3MT method in both, linear and isokinetic mode, and reported that the 3MT provided a better agreement in isokinetic mode (LoA=4 ± 30 W; SEE=5%) than in linear mode (LoA=30 ± 47 W; SEE=8%). Moreover, the authors noted significant differences and low LoA between *W'* and WEP derived from both isokinetic mode 3MT (LoA -7 ± 9 kJ; SEE 27%), and linear-mode 3MT (LoA 9±9 kJ; SEE=26%) (119).

The 'gold-standard' approach to determine CP and *W'* is still a series of CWR in the laboratory (51,60), and therefore is the method chosen to validate the 3MT (12,96,109,110). However, while several studies have reported a close agreement between traditional and 3MT derived measures of CP and EP (12,96,109,110), others have reported that EP overestimates CP, irrespective of the mathematical model used to determine CP (9,14,84). Indeed, whilst exercise at CP can be sustained for >20 min, exercise at EP was only maintained for 12–15 min (12,13,76). However, EP has demonstrated a strong positive correlation with a various thresholds, such as the lactate threshold (r = 0.79), the maximal lactate steady state (MLSS; r = 0.93), and the onset of blood lactate accumulation (r = 0.85) (100); and Black et al. (17) observed that performance in a 16.1 km cycling TT was strongly correlated with EP (r = 0.83). However, the PO associated with the MLSS was 24 W (11%) (42) to 54 W (21%) (100) lower than EP. Moreover, the difference between EP and MLSS showed heteroscedasticity, as the difference between these two parameters increased in highly trained individuals (100). Indeed, the use of the 3MT has been criticised for elite cyclists as EP overestimated CP by ~50 W, and WEP underestimated *W'* by ~8.8 kJ (9), and the difference between actual performance and the estimated performance derived from the 3MT increases with Nonetheless, 3MT is able detect changes in CP following four weeks of high-intensity training, as both CP and EP increased by a similar (r = 0.77) magnitude, and the agreement between CP and EP was good, pre- and post-training (typical error 4.6 W and 4.3 W, respectively) (111). Furthermore, Clark et al. (32) demonstrated that a 3MT is able to detect fatigue-induced changes in EP and WEP during prolonged cycling. These authors found that 2 hours of heavy exercise causes a decrease of 8% and 20% for CP and *W´*, respectively, suggesting EP and WEP may be able to assess fatigue. In summary, although 3MT may offer a time-efficient approach to estimate CP and *W'* and an ability to monitor training adaptations and fatigue, these studies suggest that a degree of caution is warranted when assuming that EP and WEP represent CP and *W'*, respectively, particularly in elite athletes.

5. Conclusions

The non-linear *P-T* relationship is well described by a hyperbolic function, which results in two parameters: the asymptote (CP), and the curvature constant (*W'*). Conventionally, several CWR to task-failure are required to determine CP and *W'*, using various modelling techniques. However, the mathematical model used, and the characteristics of the exhaustive trials such as duration, rest between trials, and mode (TT vs. CWR) have been shown to affect CP and *W'* estimations. It is recommended that CP and *W'* should be determined using the the two-parameter model that results in the lowest SEE. Regarding the exhaustive trials, a minimum of three CWR or TT is recommended with a duration spanning 2 min to 15 min. Trials which fall outside of this time range should not be used to estimate CP and *W'*, and the attainment of $\dot{V}O_{2max}$ should be verified where possible. Moreover, if the individual SEE exceeds 2-5% for CP and/or 10% for *W'*, further trials should be included in the calculation. Whilst recovery between exercise bouts of ≥60 mins appears to be sufficient to avoid $\dot{V}O_2$ priming effects, the inability to determine *W'* suggests that at present 24 h recovery periods between trials are best. The use of TT has recently been used to determine the *P-T* relationship from the field. Although there are a number of factors that might confound laboratory- vs. field-based tests, such as seating positions, acceleration and inertia, air resistance, or differences in terrain; field tests seem to provide similar CP values than those established in the laboratory whilst also offering an ecologically valid and practical approach to determine CP and *W'*. Field-based tests can be integrated into daily training, which in turn reduces the need for laboratory access and equipment. Similarly, CP testing in the laboratory can now be performed using TT. However, whilst this testing method provides highly reliable results for both parameters, it still requires further research to investigate validity of W' values. The 3MT allows the determination of EP and WEP, which are considered to represent CP and *W'*, respectively. Although a good agreement between estimates of CP and *W'* derived from the conventional approach and 3MT has been used to validate the latter; recent research suggests that EP may overestimate CP, especially in elite athletes. The original 3MT requires repeated laboratory visits: an initial GXT to determine gas exchange threshold and $\dot{V}O_{2max}$, and a subsequent visit to perform the actual 3MT. A number of alternatives have been proposed to further reduce the protocol to a single-day test. Though some of these alternatives have shown good agreement between methods, further research should also investigate the physiological responses at EP, determined from these alternatives 3MT protocols. The recommendations given in the current review should be applied to cycling, but, where possible, might be extended to other modes of exercise, such as running, swimming, rowing, or kayaking.

6. Reference List

1. Abbiss, CR and Laursen, PB. Models to explain fatigue during prolonged endurance cycling. *Sport Med* 35: 865–98, 2005.

2. Atkinson, G and Nevill, AM. Measurement Error (Reliability) in Variables Relevant to Sports Medicine. *Sport Med* 26: 217–238, 1998.

3. Bailey, SJ, Vanhatalo, A, Wilkerson, DP, Dimenna, FJ, and Jones, AM. Optimizing the ‘priming’ effect: influence of prior exercise intensity and recovery duration on O_2 uptake kinetics and severe-intensity exercise tolerance. *J Appl Physiol* 107: 1743–1756, 2009.

4. Baldwin, J, Snow, RJ, and Febbraio, MA. Effect of training status and relative exercise intensity on physiological responses in men. *Med Sci Sport Exerc 32*: 1648–54, 2000.

5. Barbosa, LF, Montagnna, L, Denadai, BS, and Greco, CC. Reliability of cardiorespiratory parameters during cycling exercise performed at the severe domain in active individuals. *J Strength Cond Res* 28: 976–981, 2014.

6. Barker, AR, Bond, B, Toman, C, Williams, C, and Armstrong, N. Critical power in adolescents: physiological bases and assessment using all-out exercise. *Eur J Appl Physiol* 112: 1359–70, 2012.

7. Barker, AR, Welsman, JR, Fulford, J, Welford, D, and Armstrong, N. Quadriceps muscle energetics during incremental exercise in children and adults. *Med Sci Sport Exerc* 42: 1303–1313, 2010.

8. Barker, T, Poole, DC, Noble, ML, and Barstow, TJ. Human critical power-oxygen uptake relationship at different pedalling frequencies. *Exp Physiol* 91: 621–632, 2006.

9. Bartram, J, Thewlis, D, Martin, D T, Norton, K I. Predicting critical power in elite cyclists: Questioning validity of the 3-min all-out test. *Int J Sports Physiol Perform* 12: 783-787, 2017.

10. Bergstrom, HC, Housh, TJ, Zuniga, JM, Camic, CL, Traylor, DA, Schmidt, RJ, et al. A new single work bout test to estimate critical power and anaerobic work capacity. *J Strength Cond Res* 26: 656–63, 2012.

11. Bergstrom, HC, Housh, TJ, Cochrane-Snyman, KC, Jenkins, NDM, Byrd, T, Switalla, JR, et al. A model for identifying intensity zones above critical velocity. *J Strength Cond Res* 31: 3260-3265, 2017.

12. Bergstrom, HC, Housh, TJ, Zuniga, JM, Traylor, D a, Lewis, RW, Camic, CL, et al. Responses during exhaustive exercise at critical power determined from the 3-min all-out test. *J Sports Sci* 31: 537–45, 2013.

13. Bergstrom, HC, Housh, TJ, Zuniga, JM, Traylor, DA, Lewis, RW, Camic, C, et al. Metabolic and neuromuscular responses at critical power from the 3-min all-out test. *Appl Physiol Nutr Metab* 38: 7–13, 2013.

14. Bergstrom, HC, Housh, TJ, Zuniga, JM, Traylor, DA, Lewis, RW, Camic, CL, et al. Differences among estimates of critical power and anaerobic work capacity derived from five mathematical models and the three-minute all-out test. *J Strength Cond Res* 28: 592–600, 2014.

15. Bishop, D and Jenkins, DG. The influence of recovery duration between periods of exercise on the critical power function. *Eur J Appl Physiol Occup Physiol* 72: 115–120,

1995.

16. Bishop, D, Jenkins, DG, and Howard, A. The critical power function is dependent on the duration of the predictive exercise tests chosen. *J Sports Med* 19: 125–9, 1998.

17. Black, MI, Durant, J, Jones, AM, and Vanhatalo, A. Critical power derived from a 3-min all-out test predicts 16.1-km road time-trial performance. *Eur J Sport Sci* 14: 217–23, 2014.

18. Black, MI, Jones, AM, Bailey, SJ, and Vanhatalo, A. Self-pacing increases critical power and improves performance during severe-intensity exercise. *Appl Physiol Nutr Metab* 40: 662–70, 2015.

19. Black, MI, Jones, AM, Blackwell, JR, Bailey, SJ, Wylie, LJ, McDonagh, STJ, et al. Muscle metabolic and neuromuscular determinants of fatigue during cycling in different exercise intensity domains. *J Appl Physiol* 122: 446–459, 2017.

20. Bosquet, L, Duchene, A, Lecot, F, Dupont, G, and Leger, L. Vmax estimate from three-parameter critical velocity models: validity and impact on 800 m running performance prediction. *Eur J Appl Physiol* 97: 34–42, 2006.

21. Brickley, G, Doust, J, and Williams, CA. Physiological responses during exercise to exhaustion at critical power. *Eur J Appl Physiol* 88: 146–151, 2002.

22. Bull, AJ, Housh, TJ, Johnson, GO, and Perry, SR. Effect of mathematical modeling on the estimation of critical power. *Med Sci Sport Exerc* 32: 526–530, 2000.

23. Bull, AJ, Housh, TJ, Johnson, GO, and Rana, SR. Physiological responses at five estimates of critical velocity. *Eur J Appl Physiol* 102: 711–20, 2008.

24. Burnley, M, Doust, JH, and Jones, AM. Time required for the restoration of normal heavy exercise $\dot{V}O_2$ kinetics following prior heavy exercise. *J Appl Physiol* 101: 1320–1327, 2006.

25. Burnley, M, Doust, JH, and Vanhatalo, A. A 3-min all-out test to determine peak oxygen uptake and the maximal steady state. *Med Sci Sport Exerc* 38: 1995–2003, 2006.

26. Busso, T, Gimenez, P, and Chatagnon, M. A comparison of modelling procedures used to estimate the power-exhaustion time relationship. *Eur J Appl Physiol* 108: 257–263, 2010.

27. Caputo, F and Denadai, BS. Exercise mode affects the time to achieve $\dot{V}O_{2max}$ without influencing maximal exercise time at the intensity associated with $\dot{V}O_{2max}$ in triathletes*. Int J Sports Med* 27: 798–803, 2006.

28. Chatagnon, M, Pouilly, J-P, Thomas, V, and Busso, T. Comparison between maximal power in the power-endurance relationship and maximal instantaneous power. *Eur J Appl Physiol* 94: 711–717, 2005.

29. Chidnok, W, Dimenna, FJ, Bailey, SJ, Wilkerson, DP, Vanhatalo, A, and Jones, AM. Effects of pacing strategy on work done above critical power during high-intensity exercise. *Med Sci Sport Exerc* 45: 1377–85, 2013.

30. Clark, IE, Gartner, HE, Williams, JL, and Pettitt, RW. Validity of the 3-minute all-out exercise test on the CompuTrainer. *J Strength Cond Res* 30: 825–829, 2016.

31. Clark, IE, Murray, SR, and Pettitt, RW. Alternative procedures for the 3-min all-out

exercise test. *J Strength Cond Res* 27: 2104-12, 2013.

32. Clark, IE, Vanhatalo, A, Bailey, SJ, Wylie, LJ, Kirby, BS, Wilkins, BW, et al. Effects of Two Hours of Heavy-Intensity Exercise on the Power–Duration Relationship. *Med Sci Sport Exerc,* 2018.

33. Constantini, K, Sabapathy, S, and Cross, TJ. A single-session testing protocol to determine critical power and W'. *Eur J Appl Physiol* 114: 1153–61, 2014.

34. Dekerle, J, Barstow, TJ, Regan, L, and Carter, H. The critical power concept in all-out isokinetic exercise. *J Sci Med Sport* 17: 640-4, 2014.

35. Dekerle, J, Brickley, G, Hammond, AJ, Pringle, JSM, and Carter, H. Validity of the two-parameter model in estimating the anaerobic work capacity. *Eur J Appl Physiol* 96: 257–64, 2006.

36. Dekerle, J, de Souza, KM, de Lucas, RD, Guglielmo, LGA, Greco, CC, and Denadai, BS. Exercise tolerance can be enhanced through a change in work rate within the severe intensity domain: work above critical power is not constant. *PLoS One* 10: e0138428, 2015.

37. Denadai, BS and Greco, CC. Can the critical power model explain the increased peak velocity/power during incremental test after concurrent strength and endurance training? *J Strength Cond Res* 31: 2319-2323, 2017.

38. Dicks, ND, Jamnick, NA, Murray, SR, and Pettitt, RW. Load determination for the 3-minute all-out exercise test for cycle ergometry. *Int J Sports Physiol Perform* 11: 197–203, 2016.

39. Ferguson, C, Rossiter, HB, Whipp, BJ, Cathcart, AJ, Murgatroyd, SR, and Ward, SA. Effect of recovery duration from prior exhaustive exercise on the parameters of the power-duration relationship. *J Appl Physiol* 108: 866–74, 2010.

40. Ferguson, C, Wilson, J, Birch, KM, and Kemi, OJ. Application of the speed-duration relationship to normalize the intensity of high-intensity interval training. *PLoS One* 8: 1–10, 2013.

41. Florence, S and Weir, JP. Relationship of critical velocity to marathon running performance. Eur *J Appl Physiol* 75: 274–278, 1997.

42. Francis, JT, Quinn, TJ, Amann, M, and Laroche, DP. Defining intensity domains from the end power of a 3-min all-out cycling test. *Med Sci Sport Exerc* 42: 1769–1775, 2010.

43. Gaesser, GA, Carnevale, TJ, Garfinkel, A, Walter, DO, and Womack, CJ. Estimation of critical power with nonlinear and linear models. *Med Sci Sport Exerc* 27: 1430–1438, 1995.

44. Gaesser, GA and Wilson, LA. Effects of continuous and interval training on the parameters of the power-endurance time relationship for high-intensity exercise. *Int J Sports Med* 9: 417–421, 1988.

45. Galbraith, A, Hopker, JG, Lelliott, S, Diddams, L, and Passfield, L. A single-visit field test of critical speed. *Int J Sports Physiol Perform* 9: 931–5, 2014.

46. Goulding, RP, Roche, DM, and Marwood, S. Prior exercise speeds pulmonary oxygen uptake kinetics and increases critical power during supine but not upright cycling. *Exp Physiol* 102: 1158-1176, 2017.

47. Hansen, EA and Smith, G. Factors affecting cadence choice during submaximal cycling and cadence influence on performance. *Int J Sports Physiol Perform* 4: 3–17, 2009.

48. Heubert, RAP, Billat, VL, Chassaing, P, Bocquet, V, Morton, RH, Koralsztein, JP, et al. Effect of a previous sprint on the parameters of the work-time to exhaustion relationship in high intensity cycling. *Int J Sports Med* 26: 583–592, 2005.

49. Hill. The relationship between power and time to fatigue in cycle ergometer exercise. *Int J Sports Med* 25: 357–61, 2004.

50. Hill, A V. The physiological basis of athletic records. *Nature* 116: 544–548, 1925.

51. Hill, DW. The critical power concept: A review. *Sport Med* 16: 273–254, 1993.

52. Hill, DW, Poole, DC, and Smith, JC. The relationship between power and the time to achieve $\dot{V}O_{2max}$. *Med Sci Sport Exerc* 34: 709–714, 2002.

53. Hinckson, E a. and Hopkins, WG. Reliability of Time to Exhaustion Analyzed with Critical-Power and Log-Log Modeling. *Med Sci Sport Exerc* 37: 696–701, 2005.

54. Hopkins, WG. Measures of reliability in sports medicine and science. *Sport Med* 30: 1–15, 2000.

55. Housh, DJ, Housh, TJ, and Bauge, SM. A methodological consideration for the determination of critical power and anaerobic work capacity. *Res Q Exerc Sport* 61: 406–409, 1990.

56. Housh, TJ, Cramer, JT, Bull, AJ, Johnson, GO, and Housh, DJ. The effect of mathematical modeling on critical velocity. *Eur J Appl Physiol* 84: 469–475, 2001.

57. Jenkins, DG, Kretek, K, and Bishop, D. The Duration of predicting trials influences time to fatigue at critical power. *J Sci Med Sport* 1: 213–218, 1998.

58. Johnson, TM, Sexton, PJ, Placek, AM, Murray, SR, and Pettitt, RW. Reliability analysis of the 3-min all-out exercise test for cycle ergometry. *Med Sci Sports Exerc* 43: 2375–2380, 2011.

59. Jones, AM and Vanhatalo, A. The ‘critical power’ concept: Applications to sports performance with a focus on intermittent high-intensity exercise. *Sport Med* 47: 65–78, 2017.

60. Jones, AM, Vanhatalo, A, Burnley, M, Morton, R, and Poole, DC. Critical power: implications for determination of $\dot{V}O_{2max}$ and exercise tolerance. *Med Sci Sport Exerc* 42: 1876–1890, 2010.

61. Jones, AM, Wilkerson, DP, Dimenna, FJ, Fulford, J, and Poole, DC. Muscle metabolic responses to exercise above and below the ‘critical power’ assessed using ^{31}P-MRS. *Am J Physiol Regul Integr Comp Physiol* 294: R585-593, 2008.

62. Jones, AM, Wilkerson, DP, Vanhatalo, A, and Burnley, M. Influence of pacing strategy on O_2 uptake and exercise tolerance. *Scand J Med Sci Sports* 18: 615–626, 2008.

63. Karsten, B, Baker, J, Naclerio, F, Klose, A, Bianco, A, and Nimmerichter, A. Time trials versus time to exhaustion tests: Effects on critical power, W’ and oxygen uptake kinetics. *Int J Sports Physiol Perform*, 2017.

64. Karsten, B, Hopker, J, Jobson, S, Baker, J, Petrigna, L, Klose, A, et al. Comparison of

inter-trial recovery times for the determination of critical power and W′ in cycling. *J Sport Sci* 35: 1420-1425, 2017.

65. Karsten, B, Jobson, SA, Hopker, J, Passfield, L, and Beedie, C. The 3-min test does not provide a valid measure of critical power using the SRM isokinetic mode. *Int J Sports Med* 35: 304–9, 2014.

66. Karsten, B, Jobson, SA, Hopker, J, Jimenez, A, and Beedie, C. High agreement between laboratory and field estimates of critical power in cycling. *Int J Sports Med* 35: 298–303, 2014.

67. Karsten, B, Jobson, SA, and Hopker, JG. Validity and reliability of critical power field testing. *Eur J Appl Physiol* 115: 197–204, 2015.

68. Kennedy, M and Bell, DG. A comparison of critical velocity estimates to actual velocities in predicting simulated rowing performance. *Can J Appl Physiol* 25: 223–235, 2000.

69. Kennely. An approximate law of fatigue in the speeds of racing animals. *Proc Am Acad Arts Sci* 42: 275–331, 1906.

70. Kranenburg KJ and Smith, DJ. Comparison of critical speed determined from track running and treadmill tests in elite runners. *Med Sci Sport Exerc* 28: 614–8, 1999.

71. Lansley, KE, DiMenna, FJ, and Jones, AM. A ‘new’ method to normalise exercise intensity. *Int J Sports Med* 32: 535–41, 2011.

72. Laursen, PB, Francis, GT, Abbiss, CR, Newton, MJ, and Nosaka, K. Reliability of time-to-exhaustion versus time-trial running tests in runners. *Med Sci Sports Exerc* 39: 1374–9, 2007.

73. de Lucas, RD, Greco, CC, Dekerle, J, Caritá, RA, Guglielmo, LGA, and Denadai, BS. Test-retest reliability of a 3-min isokinetic all-out test using two different cadences. *J Sci Med Sport* 17: 645–649, 2014.

74. Mann, T, Lamberts, RP, and Lambert, MI. Methods of prescribing relative exercise intensity: Physiological and practical considerations. *Sport Med* 43: 613–625, 2013.

75. Mattioni Maturana, F, Fontana, FY, Pogliaghi, S, Passfield, L, and Murias, JM. Critical power: How different protocols and models affect its determination. *J Sci Med Sport* 1–6, 2017.

76. McClave, SA, LeBlanc, M, and Hawkins, SA. Sustainability of critical power determined by a 3-minute all-out test in elite cyclists. *J Strength Cond Res* 25: 3093–3098, 2011.

77. McLellan, TM and Cheung, KS. A comparative evaluation of the individual anaerobic threshold and the critical power. *Med Sci Sports Exerc* 24: 543–550, 1992.

78. Morin, JB and Sève, P. Sprint running performance: Comparison between treadmill and field conditions. *Eur J Appl Physiol* 111: 1695–1703, 2011.

79. Moritani, T, Nagata, A, DeVries, HA, and Muro, M. Critical power as a measure of physical work capacity and anaerobic threshold. *Ergonomics* 24: 339–50, 1981.

80. Morton, RH. A 3-parameter critical power model. *Ergonomics* 34: 611–9, 1996.

81. Morton, RH. The critical power and related whole-body bioenergetic models. *Eur J Appl Physiol* 96: 339–354, 2006.

82. Muniz-Pumares, D, Pedlar, CR, Godfrey, R, and Glaister, M. Accumulated oxygen deficit during exercise to exhaustion determined at different supramaximal work-rates. *Int J Sports Physiol Perform* 12: 351-356, 2017.

83. Murgatroyd, SR, Ferguson, C, Ward, S a, Whipp, BJ, and Rossiter, HB. Pulmonary O_2 uptake kinetics as a determinant of high-intensity exercise tolerance in humans. *J Appl Physiol* 110: 1598–606, 2011.

84. Nicolò, A, Bazzucchi, I, and Sacchetti, M. Parameters of the 3-minute all-out test: Overestimation of competitive-cyclist time-trial performance in the severe-intensity domain. *Int J Sports Physiol Perform* 12: 655–661, 2017.

85. Nimmerichter, A, Novak, N, Triska, C, Prinz, B, and Breese, BC. Validity of treadmill-derived critical speed on predicting 5000-meter track-running performance. *J Strength Cond Res* 31: 706-714, 2017.

86. Passfield, L, Hopker, J, Jobson, S, Friel, D, and Zabala, M. Knowledge is power: Issues of measuring training and performance in cycling. *J Sports Sci* 35: 1426–1434, 2017.

87. Pepper, ML, Housh, TJ, and Johnson, GO. The accuracy of the critical velocity test for predicting time to exhaustion during treadmill running. *Int J Sports Med* 13: 121–4, 1992.

88. Pettitt, RW, Jamnick, N, and Clark, IE. 3-min All-out Exercise Test for Running. *Int J Sports Med* 33: 426-31, 2012.

89. Pinot, J and Grappe, F. The record power profile to assess performance in elite cyclists. *Int J Sports Med* 32: 839–844, 2011.

90. Poole, DC, Burnley, M, Vanhatalo, A, Rossiter, HB, and Jones, AM. Critical power: An important fatigue threshold in exercise physiology. *Med Sci Sport Exerc* 48: 2320–2334, 2016.

91. Poole, DC, Ward, SA, Gardner, GW, and Whipp, BJ. Metabolic and respiratory profile of the upper limit for prolonged exercise in man. *Ergonomics* 31: 1265–1279, 1988.

92. di Prampero, PE. The concept of critical velocity: a brief analysis. *Eur J Appl Physiol* 80: 162–164, 1999.

93. Puchowicz, MJ, Mizelman, E, Yogev, A, Koehle, MS, Townsend, NE, and Clarke, DC. The Critical Power Model as a Potential Tool for Anti-doping. *Front Physiol* 9: 643, 2018

94. Sawyer, BJ, Morton, RH, Womack, CJ, and Gaesser, G a. $\dot{V}O_{2max}$ may not be reached during exercise to exhaustion above critical power. *Med Sci Sports Exerc* 44: 1533–8, 2012.

95. Scherrer, J and Monod, H. Local muscle work and fatigue in man. *J Physiol* 52: 419–501, 1960.

96. Simpson, LP, Jones, AM, Skiba, PF, Vanhatalo, A, Wilkerson, D, Sciences, H, et al. Influence of hypoxia on the power-duration relationship during high-intensity exercise. *Int J Sports Med* 36: 113–9, 2015.

97. Simpson, LP and Kordi, M. Comparison of critical power and W' derived from two or three maximal tests. *Int J Sports Physiol Perform* 12: 825-830, 2017.

98. Skiba, PF, Chidnok, W, Vanhatalo, A, and Jones, AM. Modeling the expenditure and reconstitution of work capacity above critical power. *Med Sci Sports Exerc* 44: 1526–

1532, 2012.

99. Smith, JC, Dangelmaier, BS, and Hill, DW. Critical power is related to cycling time trial performance. *Int J Sports Med* 20: 374–378, 1999.

100. Sperlich, B, Haegele, M, Thissen, A, Mester, J, and Holmberg, HC. Are peak oxygen uptake and power output at maximal lactate steady state obtained from a 3-min all-out cycle test? *Int J Sports Med* 32: 433–437, 2011.

101. Townsend, NE, Nichols, DS, Skiba, PF, Racinais, S, and Périard, JD. Prediction of critical power and *W′* in hypoxia: Application to work-balance modelling. *Front. Physiol.* 8: 180, 2017.

102. Triska, C, Karsten, B, Beedie, C, Koller-zeisler, B, Nimmerichter, A, Tschan, H, et al. Different durations within the method of best practice affect the parameters of the speed – duration relationship. *Eur J Sport Sci.* 2018.

103. Triska, C, Karsten, B, Heidegger, B, Koller-Zeisler, B, Prinz, B, Nimmerichter, A, et al. Reliability of the parameters of the power-duration relationship using maximal effort time-trials under laboratory conditions. *PLoS One* 12: e0189776, 2017.

104. Triska, C, Karsten, B, Nimmerichter, A, and Tschan, H. Iso-duration determination of D′ and CS under laboratory and field conditions. *Int J Sports Med* 38: 527–533, 2017.

105. Triska, C, Tschan, H, Tazreiter, G, and Nimmerichter, A. Critical power in laboratory and field conditions using single-visit maximal effort trials. *Int J Sports Med* 36: 1063–1068, 2015.

106. Vandewalle, H, Pérès, G, and Monod, H. Standard anaerobic exercise tests. *Sport Med* 4: 268–289, 1987.

107. Vandewalle, H, Vautier, JF, Kachouri, M, Lechevalier, JM, and Monod, H. Work-exhaustion time relationships and the critical power concept: A critical review. *J Sports Med Phys Fitness* 37: 89–102, 1997.

108. Vanhatalo, A, Black, MI, Dimenna, FJ, Blackwell, JR, Schmidt, JF, Thompson, C, et al. The mechanistic bases of the power-time relationship: muscle metabolic responses and relationships to muscle fibre type. *J Physiol* 594: 4407–23, 2016.

109. Vanhatalo, A, Doust, JH, and Burnley, M. Determination of critical power using a 3-min all-out cycling test. *Med Sci Sports Exerc* 39: 548–55, 2007.

110. Vanhatalo, A, Doust, JH, and Burnley, M. Robustness of a 3 min all-out cycling test to manipulations of power profile and cadence in humans. *Exp Physiol* 93: 383–390, 2008.

111. Vanhatalo, A, Doust, JH, and Burnley, M. A 3-min all-out cycling test is sensitive to a change in critical power. *Med Sci Sports Exerc* 40: 1693–9, 2008.

112. Vanhatalo, A, Fulford, J, DiMenna, FJ, and Jones, AM. Influence of hyperoxia on muscle metabolic responses and the power-duration relationship during severe-intensity exercise in humans: a 31P magnetic resonance spectroscopy study. *Exp Physiol* 95: 528–540, 2010.

113. Vanhatalo, A, McNaughton, LR, Siegler, J, and Jones, AM. Effect of induced alkalosis on the power-duration relationship of ‘all-out’ exercise. *Med Sci Sport Exerc* 42: 563–70, 2010.

114. Vanhatalo, A, Poole, DC, DiMenna, FJ, Bailey, SJ, and Jones, AM. Muscle fiber recruitment and the slow component of O_2 uptake: constant work rate vs. all-out sprint exercise. *Am J Physiol Regul Integr Comp Physiol* 300: R700-7, 2011.

115. Whipp, BJ and Ward, SA. Quantifying intervention-related improvements improvements in exercise tolerance. *Eur Respir J* 33: 1254–1260, 2009.

116. Wilkie, DR. Equations describing power input by humans as a function of duration of exercise. In: Exercise bioenergetics and gas exchange. Cerretelli, P and Whipp, BJ, eds. Amsterdam: Elsevier/North Holland Biomedical Press, 1980. pp. 74–80

117. Winter, E, Abt, G, Brookes, F, Challis, J, Fowler, N, Knudson, D, et al. Misuse of ‘Power’ and other mechanical terms in sport and exercise science research. *J Strength Cond Res* 30: 292–300, 2016.

118. Wright, J, Jobson, S, and Bruce-Low, S. The reliability and validity of the 3-minute critical power test. *Int J Sports Med* 38: 462–467, 2017.

7. Tables and Figures

Table 1. Example of data collected from five constant-work rate bouts to task-failure in a trained cyclist. Power and Duration are recorded during the test, and work and Time^{-1} subsequently calculated. ‘Max’ represents peak power output.

Trial	Power (W)	Duration (s)	Work (kJ)	Time^{-1} (s^{-1})
1	415	135	56.03	0.0074
2	360	240	86.40	0.0042
3	340	408	138.72	0.0025
4	320	600	192.00	0.0017
5	310	930	288.30	0.0011
Max	1100			

Figure Legends

Figure 1. Different modelling approaches to determine critical power and the curvature constant *W’* from data presented in Table 1. Panel A represents the 2-parameter hyperbolic power-duration relationship. Panel B represents the 3-parameter hyperbolic power-duration relationship. Panel C represents the 2-paremeter linear work-T_{lim} relationship. Panel D represents the 2-parameter linear power output- T_{lim}^{-1} relationship. T_{lim} represent duration until task-failure.

Figure 2. The effect of the different mathematical modelling approaches to determine critical power and *W’* on the relationship between power output and time to task-failure. Data from Table 1.

Figure 3. The effect of the duration of the trial on critical power (CP) and *W’*. Data from Table 1.

Figure 4. Outline of the 3-min all-out test. Panel A represents data from 30 seconds before the start of the test (start at time = 0 s). Panel B represents 30-seconds averages through the test. Filled circles (●) denote power output, and open circles (○) represent oxygen consumption ($\dot{V}O_2$). Note that power output initially increases, reaching a peak in the first few seconds of the test, and then progressively decreases until, eventually levels off in the final 30 s of the test (i.e. end-test power output).