# Different durations within the method of best practice affect the parameters of the speed-duration relationship

The aim of the study was to determine whether estimates of the speed-duration relationship are affected using different time-trial (TT) field-based testing protocols, where exhaustive times were located within the generally recommended durations of 2 to 15 min. Ten triathletes (mean±SD age: 31.0±5.7yrs; height: 1.81±0.05m; body mass: 76.5±6.8kg) performed two randomly assigned field-tests to determine critical speed (CS) and the total distance covered above CS (*D´*). CS and *D´* were obtained using two different protocols comprising three TT that were interspersed by 60 min passive rest. The TTs were 12, 7, and 3 min in *Protocol I* and 10, 5, and 2 min in *Protocol II*. A linear relationship of speed vs. the inverse of time (s=*D´*x1/t+CS) was used to determine parameter estimates. Significant differences were found for CS (*P*=.026), but not for *D´* (*P*=.123). The effect size for CS (*d*=.305) was considered small, whilst that for *D´* was considered moderate (*d*=.742). CS was significantly correlated between protocols (*r*=.934; *P*<.001), however, no correlation was found for *D´* (*r*=.053; *P*=.884). The 95% limits of agreement were ±0.28m·s$^{-1}$ and ±73.9m for CS and *D´*, respectively. These findings demonstrate that the choice of exhaustive times within commonly accepted durations, results in different estimates of CS and *D´* and thus protocols cannot be used interchangeably. The use of a consistent protocol is therefore recommended, when investigating or monitoring the speed-duration relationship estimates in well-trained athletes.

Keywords: maximum effort running, performance, exhaustion, field testing

## Introduction

A linear relationship between speed and the inverse of time in running was first demonstrated by Hughson, Orok, and Staudt (1984). The parameter estimates of this relationship serve as important parameters for performance assessment (e.g. Hughson et al., 1984; Jones, Vanhatalo, Burnley, Morton, & Poole, 2010), training prescription (e.g. Galbraith, Hopker, & Passfield, 2015), as well as performance prediction (e.g. Florence

& Weir, 1997; Kranenburg & Smith, 1996; Nimmerichter, Novak, Triska, Prinz, & Breese, 2017). Critical speed (CS) functions as a demarcation line between the heavy and the severe exercise intensity domain, whilst its related total distance covered above CS (*D´*) serves as an important parameter for high-intensity exercise (Bull, Housh, Johnson, & Rana, 2008; Jones et al., 2010). Knowing these two parameters the tolerance of high intensity exercise >CS can be predicted accurately (Ferguson, Wilson, Birch, & Kemi, 2013).

The currently recommended methods of best-practice of speed-duration relationship testing generally involves repeated exhaustive trial durations ($t_{lim}$) between 2 and 15 min, with a minimum of 5-min difference between the shortest and the longest trial (Galbraith, Hopker, Jobson, & Passfield, 2011; Hill, Vingren, Nakamura, & Kokobun, 2011; Housh et al., 1991; Karsten et al., 2016; Nimmerichter et al., 2017; Triska, Karsten, Nimmerichter, & Tschan, 2017). However, there is no consensus as to which value of $t_{lim}$ produces the least error containing, i.e. the most valid parameter estimates (Busso, Gimenez, & Chatagnon, 2010; Vandewalle, Vautier, Kachouri, Lechevalier, & Monod, 1997). Whilst some authors utilised 2 min for the shortest trial (Karsten et al., 2016; Nimmerichter et al., 2017) others utilised 3 min (Bergstrom et al., 2017; Galbraith et al., 2011; Hill et al., 2011; Triska et al., 2017). Similarly, durations of the longest trials vary between 10 and 15 min (Galbraith et al., 2011; Hill et al., 2011; Karsten et al., 2016; Nimmerichter et al., 2017; Triska et al., 2017). This demonstrates a current lack of knowledge if different $t_{lim}$-protocols within a recommended time range can be used interchangeably to validly estimate CS and *D´*.

Previous research focusing on the effect of different $t_{lim}$ on the parameter estimates have had two major limitations: Bishop, Jenkins, and Howard (1998) and Busso et al. (2010) compared protocols that did not follow the recommendation of a 5-

min minimum time difference between the longest and the shortest trial, whilst Jenkins, Kretek, and Bishop (1998) used $t_{lim}$ outside the time-band between 2 and 15 min that is commonly used in recent studies. To date, research has not yet addressed the question whether $t_{lim}$ differences within these accepted limits affect the parameter estimates when complying with the previously proposed requirement of a time difference of >5 min between the longest and the shortest trial.

Attaining $\dot{V}O_{2max}$ and discharging *D´* at the end of each exhaustive run trial is a pre-requisite of an accurate determination of the speed-duration relationship (di Prampero, 1999). As $\dot{V}O_{2max}$ might not be attained and *D´* might not be fully depleted, $t_{lim}$ <2 min should consequently be avoided (di Prampero, 1999). Conversely, due to a too low intensity even during exercise in the severe intensity domain, but also due to motivational factors, $\dot{V}O_{2max}$ might not be attained with $t_{lim}$ >15 min (Poole, Ward, Gardner, & Whipp, 1988; Sawyer, Morton, Womack, & Gaesser, 2012; Vandewalle et al., 1997). For example, with lower levels of muscle activation and metabolic demands, Bergstrom et al. (2017) reported that $\dot{V}O_{2max}$ was not attained in $t_{lim}$ >12 min. Vandewalle et al. (1997) indicated that $t_{lim}$ below 3 min and above 30 min deviate from the regression line resulting in altered values of CS and *D´*. Whilst shorter trials (~1 min) generally estimate CS higher and *D´* lower, longer trials (~30 min) in turn generally estimate CS lower and *D´* higher (Vandewalle et al., 1997).

Recently, questions have been raised about a valid and reliable estimation of *D´*. For example, Galbraith et al. (2015) demonstrated a significant difference between actual *D´* and estimated *D´* using a single-visit field test. Moreover, *D´* has not yet been translated validly (Galbraith, Hopker, Lelliott, Diddams, & Passfield, 2014; Karsten, Jobson, Hopker, Jimenez, & Beedie, 2014; Triska et al., 2017; Triska, Tschan, Tazreiter, & Nimmerichter, 2015) or reliably (Galbraith et al., 2011; Karsten, Jobson,

Hopker, Stevens, & Beedie, 2015) from the laboratory into the field. Respective discussions either focussed on a high variability for this parameter (>80 m in Galbraith et al. (2014)). Importantly, in their criterion (i.e. time-to-exhaustion; TTE) and experimental (i.e. TT) protocols, relevant studies consistently used different $t_{lim}$, resulting in high levels of agreement for CS but not for *D´* (Galbraith et al., 2014; Galbraith et al., 2011; Hill et al., 2011; Karsten et al., 2016; Nimmerichter et al., 2017; Triska et al., 2017). This was recently suggested to account for some of the differences by Triska et al. (2017) who identified no significant differences but also significant correlations for values of *D´* when using equal $t_{lim}$ across protocols and respective runs.

The novelty of this study was therefore to compare two protocols that are commonly used in research comprising $t_{lim}$ between 2 and 12 min with a time difference of >5 min between the longest and shortest trial. A further aim was to assess potential differences in CS and *D´* when comparing two different single-visit protocols. As the speed-duration relationship shows a linearity within the recommended time range (Hill, 1993), we hypothesized non-significant differences for CS and *D´* and a high level of agreement (i.e. a small bias and a significant correlation) for CS and *D´* between the protocols.

## Methods

### *Ethical Approval*

All procedures were submitted to and approved by the host institution's Ethics Committee (Ref. Nbr. 00155) and conformed to the principles of the *Declaration of Helsinki*. Participants completed a health questionnaire and provided written informed consent after being fully informed about all experimental procedures.

***Participants***

Ten endurance trained male triathletes (mean ± SD age: 31.0 ± 5.7 yrs; height: 1.81 ± 0.05 m; body mass: 76.5 ± 6.8 kg; maximal aerobic speed (MAS): 4.59 ± 0.31 $m \cdot s^{-1}$) performed two different performance tests. Participants trained for approximately 8 hours per week, were familiar with TT runs, and had at least 3-years experience in running competitions and triathlons at a national and international level. They were instructed to avoid strenuous exercises for 24 h prior to each testing session, to abstain from alcohol and caffeine on the day of testing, and to arrive for all tests 3 hours postprandial in a fully hydrated state.

***Design***

The procedures followed a repeated field test design after a preliminary GXT to determine MAS. To determine CS and *D´*, field tests comprised of three exhaustive TT runs of different durations on a 400-m athletic outdoor track. To allow runners to follow the line of the least distance (i.e. 400 m), the track during all tests was only used by our participants. It was ensured that the participant were able to follow the line of least distance. The exhaustive runs were conducted in the order from the longest to the shortest duration (e.g. Galbraith et al., 2011; Jenkins & Quigley, 1992; Triska et al., 2017) and were interspersed by 60 min passive rest (Karsten, Baker, et al., 2017), where participants were allowed to drink water ad libitum. Single-visit methods (i.e. 3 maximal effort runs or a single all-out run) are commonly used in current literature (e.g. Broxterman, Ade, Poole, Harms, & Barstow, 2013; Galbraith et al., 2015; Galbraith et al., 2011; Karsten, Baker, et al., 2017; Karsten et al., 2016; Nimmerichter et al., 2017; Triska et al., 2017; Triska et al., 2015) and has shown to provide valid results (Galbraith et al., 2014; Karsten, Hopker, et al., 2017; Karsten et al., 2014; Karsten et al., 2015; Triska et al., 2015). The recovery interval was chosen to allow blood lactate

concentration ([La]) to return to resting values as elevated [La] has shown to alter performance in a subsequent trial (Burnley, Davison, & Baker, 2011; Burnley, Doust, & Jones, 2005). Blood samples for the determination of [La] were taken before and after each maximal run. Using an ANT+ heart rate monitor heart rate (HR) was measured thoughout all trials. The rationale for choosing the single-visit method was to compare tests that are used in a research as well as in practical settings. Runs were only performed with wind speeds <3 $m \cdot s^{-1}$ obtained from a wind gauge placed next to the track (Triska et al., 2017). The second field test was only conducted when weather conditions were similar to the first test and wind speed within limits (± 3 $m \cdot s^{-1}$ and ± 5° C for wind speed and temperature, respectively). Temperature and humidity were within a range between 5°C and 20°C and between 30% and 55%, respectively. Participants were strongly verbally encouraged during all exhaustive trials and were instructed to use the same running shoes for all tests. All testing sessions were separated by at least 72 hours and were completed within a two-week period. Tests were completed at the same time of the day (± 2 h).

***Incremental exercise test***

To determine MAS, an incremental treadmill test (Saturn, h/p cosmos Sport and Medical, Traunstein, Germany) was performed prior to formal data collection. The test commenced at 2.22 $m \cdot s^{-1}$ and speed was increased by 0.28 $m \cdot s^{-1}$ every three minutes until volitional exhaustion. If the last work stage could not be fully completed, MAS was estimated using following equation:

$$MAS = sL + t/180 \times 0.14 \quad (1)$$

Where MAS is the maximal aerobic speed, sL is the speed of the last fully completed stage ($m \cdot s^{-1}$), and t is the time of the incomplete stage (s).

***Blood lactate sampling***

To determine [La] 20 µl blood samples from a hyperemic earlobe were collected. Immediately after taking, samples were diluted in a 1000 µl glucose system solution. After that, samples were analysed by a lactate analyser (Biosen S_line, EKF Diagnostics, Barleben, Germany).

***Protocol I***

The parameters of the speed-duration relationship were determined from three maximum effort TT of different durations, which were 12, 7, and 3 min. Similar to previous research, distance was measured to the nearest metre (Triska et al., 2017). Time was measured using a running watch (Forerunner 235, Garmin International Inc. Kansas, USA), where the remaining time of the respective trial was displayed. Unpublished observations from our laboratory have shown that mean distance estimated using the running watch on a 400-m lap was ~3.3% greater compared to actual distance (range: 407 to 421 m). Due to this reduced accuracy of GPS/GLONASS, participants were consequently blinded for speed during the runs and therefore total distance covered was not taken from the watch.

Prior to each TT, participants performed a 5-min self-paced low-intensity warm-up exercise followed by 5 min stretching exercise (Galbraith et al., 2014). Timing started with a transition from walking to running and participants were instructed to cover the greatest distance possible in the set time.

Blood samples were taken ~3 min before each run and immediately (within the first minute), 3, 6, and 9 min after the end of the exhaustive. [$La_{max}$] was taken as the highest value across all samples.

***Protocol II***

*Protocol II* was conducted similar to *Protocol I* with the TT runs performed over 10, 5, and 2 min. To minimize negative effects of learning, protocols were employed in a randomised order. Warm-up procedures and blood sampling were similar to *Protocol I*.

***Determination of CS and D´***

CS and *D´* were determined using a linear regression analysis of speed and the inverse of time (Equation 2), where speed is in $m \cdot s^{-1}$, *D´* is the total distance covered above CS until task failure (m), 1/t is the inverse of time ($s^{-1}$), and CS is the critical speed ($m \cdot s^{-1}$):

$$\text{speed} = D' \times 1/t + CS \quad (2)$$

*D´* is represented by the slope, and CS by the y-intercept. The speed was defined as mean speed during a trial which was calculated as a quotient of distance and time. This model was chosen as compared to other models, it has shown to provide the lowest SE for both parameter estimates and but also, as it can easily be used in an applied setting (Nimmerichter, Steindl, & Williams, 2015). The SE of both parameter estimates was computed for each participant for absolute and relative values. SE of CS and *D´* were required to fall below 2% and 10%, respectively (Dekerle et al., 2015; Ferguson et al., 2013) and if violated, trials had to be repeated on another occasion. This was required for one participant.

***Statistical analyses***

All data were assessed using the Shapiro-Wilk test for normality. A paired samples t-test assessed differences in CS and *D´* between the protocols. Effect size was calculated using Cohen's *d* (small $d = 0.2$; moderate $d = 0.5$; large $d = 0.8$). A two-way ANOVA was used to assess differences in pre- and post [La] between trials and protocols. Significant main effects were followed-up by Tukey's post-hoc procedures. Pearson

moment product and the standard error of the estimate (SEE) assessed the relationship between the protocols. Agreement between the protocols was evaluated using 95% limits of agreement (LoA) (Bland & Altman, 1986). An alpha level of $P < .050$ was considered to be statistically significant and results are reported as mean ± SD. The smallest worthwhile effect was assumed to be 15 m for $D´$ and 0.15 $m \cdot s^{-1}$ for CS (Triska et al., 2017). An *a priori* power analysis was performed and revealed that 10 participants were required to detect a difference at an alpha-level of $P < .050$ with a statistical power of 80% (Faul, Erdfelder, Buchner, & Lang, 2009).

**Results**

Results for CS and $D´$ are shown in Table 1 and Figure 2. Between protocols, $D´$ was not significantly different ($t_9 = 1.704$; $P = .123$; $d = .742$), but a significant difference was found for CS ($t_9 = 2.654$; $P = .026$; $d = .305$). Significant correlations were found for CS ($r = .934$; $P < .001$; Figure 1a), but not for $D´$ ($r = .053$; $P = .884$; Figure 1b). The mean bias for CS was 0.12 ± 0.14 $m \cdot s^{-1}$ (95% LoA: -0.40 to 0.16 $m \cdot s^{-1}$) and for $D´$ it was 20.3 ± 37.7 m (95% LoA: -53.6 m to 94.2 m) (Table 1) (Figure 1c and 1d). Non-significant differences were found for the relative SE ($t_9 = .802$; $P = .802$; $d = .140$ and $t_9 = .481$; $P = .642$; $d = .223$ for CS and $D´$, respectively) and absolute SE ($t_9 = .330$; $P = .749$; $d = .182$ and $t_9 = .801$; $P = .444$; $d = .417$ for CS and $D´$, respectively) between protocols. Mean [La] values and maximal heart rate ($HR_{max}$) for the exhaustive runs are shown in Table 2. A two-way ANOVA revealed a significant main effect for post [La] between the trials ($F_{2,54} = 4.998$; $P = .010$), with no significant post-hoc differences ($P = .180$-$1.000$). Furthermore, no significant differences were found between the protocols ($F_{1,54} = .407$; $P = .526$) nor any interactions trial x protocol ($F_{2,54} = .020$; $P = .981$) for post [La]. No significant differences were found for pre [La] between trials ($F_{2,54} = 2.835$; $P = .068$), protocols ($F_{1,54} = .010$; $P = .917$) nor any interactions trial x protocol

($F_{2,54}$ = .067; $P$ = .935). No significant main effects (trial, protocol, or trial x protocol) for $HR_{max}$ were found ($P$ = .570-.953).

**Discussion**

This is the first study which demonstrates that the use of different protocols comprising of $t_{lim}$, which are located within the currently recommended testing method of best practice, affects estimates of CS and *D´*. This is important as these effects occurred in a cohort of well-trained athletes who produce lower levels of biological variability (Hopkins & Hewson, 2001).

The effect size of *D´* is of a moderate order despite the non-significant differences observed ($d$ = .742). This might have been the cause of a possible *type II error* as a result of the low statistical power (Buchheit, 2016). However, the small effect size observed for CS ($d$ = .305) leads to significantly faster predicted 5-km performance times using *Protocol II* parameter estimates (mean difference: 30.8 ± 37.5 s; $P$ = .029). The raw difference of CS between protocols was slightly below and the raw difference for *D´* between protocols was above than the estimated smallest worthwhile effect. Therefore, statistical power might be lower than expected for estimates of CS and effect sizes should be considered (Buchheit, 2016). A mean difference of ~3% between the protocols in CS seems very close. However, (Galbraith et al., 2011) in well-trained runners reported a coefficient of variation (CoV) of ~1.3% for CS. Therefore, it is not unreasonable to argue that these significant differences between the protocols can be interpreted as a physiological meaningful difference. This is also true for *D´*, which demonstrated a ~15% difference in the present study, whilst a CoV of ~9.8% was reported by (Galbraith et al., 2011).

Our findings are supported by previous works, i.e. that shorter trials tend

towards lower estimates of *D´* and towards higher estimates of CS (Bishop et al., 1998; Busso et al., 2010; Jenkins et al., 1998; Vandewalle et al., 1997). However, the novel findings of this study were that these differences were also evident when using protocols within the currently recommended testing method of best practice (i.e. $t_{lim}$ between two and 12 min with a >5 min difference between longest and shortest trial). Whilst CS under *Protocol II* condition demonstrated this tendency, values of *D´* however were not generally lower compared to *Protocol I*. Moreover, *D´* as in other studies (Galbraith et al., 2014; Galbraith et al., 2011; Karsten, Hopker, et al., 2017) showed a high within-subject variability and a non-significant correlation between protocols ($r$ = .053) (Figure 1b). It is unlikely that learning effects are responsible for the differences in the parameter estimates as, had they been present, the randomised design would have resulted in an even distribution of such effect between the protocols.

di Prampero (1999) recommended to avoid trials <2 min, and Vandewalle et al. (1997) and Poole et al. (1988) stated to avoid trials >12 min. Moreover, Vandewalle et al. (1997) stipulated that trials <3 min deviate from the regression line altering the parameter estimates. Therefore, $t_{lim}$ runs <3 min or >12 min would likely increase the SE for both parameter estimates due to the non-linearity of the speed-duration relationship reported by (Vandewalle et al., 1997). However, in the present study SE values were well within accepted limits proposed (Dekerle et al., 2015; Ferguson et al., 2013) and they were not significantly different between protocols, even though $t_{lim}$ <3 min were used in *Protocol II* ($P$ = .642-.802) (Table 1). Furthermore, non-significantly differences between efforts were found for post trial blood lactate concentrations [La] ($P$ = .309-.100) or $HR_{max}$ values ($P$ = .570-.953) (Table 2). Therefore, it is assumed that neither motivational aspects nor an incomplete discharge of *D´* (i.e. no exhaustion at the end of the runs) have affected the speed-duration relationship.

Interestingly, the raw difference in CS values between protocols demonstrated to be larger for runners with a lower CS (i.e. <4.0 $m \cdot s^{-1}$) compared to faster runners (Figure 1a). With the largest deviation from the regression line, the longer chosen $t_{lim}$ in *Protocol I* (in particular the 12-min run) appears to have negatively affected results in participants with a CS <4.0 $m \cdot s^{-1}$ (Figure 2). Physiologically, it could be speculated that *D´* in these participants was not fully depleted during the 12-min run despite non-significant different end test [La] and $HR_{max}$ trial values. Therefore, it is suggested that the 12-min run might have reduced the validity of the speed-duration relationship under *Protocol I* conditions for some participants. Due to the significantly higher raw difference in CS between the protocols, participants with CS <4.0 $m \cdot s^{-1}$ (n = 3) were omitted from further analysis. Interestingly, this further analysis revealed no significant differences between protocols for CS and *D´* and a trivial and small effect size for both parameter estimates (*P* = .205; *d* = .032 and *P* = .684; *d* = .213 for CS and *D´* respectively). Furthermore, a strong relationship for CS was found between the two protocols (*r* = .912; *P* = .004), however, such a trend could not be observed for *D´* and results again show a high within-subject variability (*r* = .130; *P* = .782). Moreover, this demonstrates that the mean bias between the protocols for CS and *D´* was lower compared to the original analysis comprising of the whole participation group (-0.06 ± 0.11 $m \cdot s^{-1}$ vs -0.25 ± 0.11 $m \cdot s^{-1}$ and 5.3 ± 33.0 m vs 55.3 ± 22.7 m for CS and *D´* respectively). However, the agreement for *D´* was consistently low.

Figure 2 also shows the 7-min run (i.e. medium intensity run of *Protocol I*) below the regression line of *Protocol II*, whilst the 3-min run (black arrow in Figure 2) fits the regression line of *Protocol II*. Even though not suggested as influencing CS values (Karsten, Hopker, et al., 2017), it might be that residual fatigue of the 12-min run affected performances in the 7-min run despite a longer recovery (i.e. 60 min) compared

to other works (e.g. Galbraith et al., 2011; Karsten, Hopker, et al., 2017; Triska et al., 2017) and despite [La] having returned close to resting values. Our data also revealed that mean speed between the 7-min trial (second trial in *Protocol I*) and the 10-min trial (first trial in *Protocol II*) was not significantly different ($P = .345$), whilst significant differences were found between all other trials ($P < .036$). Importantly, according to Ferguson et al. (2010) *D´* is fully reconstituted within 20 to 25 min and therefore should have been fully replenished at the onset of each consecutive $t_{lim}$ run. A number of researchers demonstrated that estimates derived from a single-visit protocol using 30 min or 60 min passive rest between efforts does not results in significant differences for CS (Galbraith et al., 2014; Karsten, Hopker, et al., 2017). However, estimates of *D´* were significantly different between inter-trial recovery protocols (Galbraith et al., 2014; Karsten, Hopker, et al., 2017). Therefore, differences might have been caused by either, physiological or alternatively psychological residual fatigue (for review see: Van Cutsem et al. (2017)) induced by the previous run as compared to CS, *D´* appears to be more sensitive to previous maximal efforts.

A single 3-min all-out trial could potentially alleviate these negative effects of residual fatigue (Broxterman et al., 2013). However, our further analysis showed a high intra-individual variation in CS and *D´* (CoV = 10.2% and 41.1% respectively) and wide limits of agreement (±0.62 $m \cdot s^{-1}$ and ±157.5 m). Moreover whilst validating CS, Broxterman et al. (2013) cautioned the use of an all-out 3 min test for the determination of *D´*.

Consistent with recent results, it appears that *D´* is either sensitive to changes in $t_{lim}$ (Triska et al., 2017), cannot be determined accurately using the speed-duration relationship (Karsten, Hopker, et al., 2017; Triska et al., 2017), or is associated with high day-to-day variability (Galbraith et al., 2011; Hinckson & Hopkins, 2005; Karsten,

Hopker, et al., 2017; Triska et al., 2017). For example Triska et al. (2017) demonstrated that matching durations (i.e. iso-duration testing) of similar efforts resulted in a significant correlation for *D´*, supporting the suggestion that *D´* is highly sensitive to changes in $t_{lim.}$ However, the authors also reported a high typical error (39.2 m or 18.7%) which suggests that the determination of *D´* is influenced also by additional factors such as significant different distances covered in one of the respective runs. Similar to our findings, Galbraith et al. (2015) also demonstrated a high variability for *D´* and the authors suggested that *D´* might be subject to day-to-day variations. Moreover, if *D´* were solely dependent on $t_{lim}$, then a significant correlation in *D´* similar to CS should have been evident in the present study (*r* =.053; Figure 1b).

**Limitations of the study**

No continuous measurement of $\dot{V}O_2$ was conducted to assess $\dot{V}O_{2max}$ and $\dot{V}O_2$ on-kinetics during the runs. Therefore, attainment of $\dot{V}O_{2max}$ at the end of the trials could not be verified and it could only be speculated that $\dot{V}O_{2max}$ was not attained in some participants during the predictive runs. The potential presence of "primed" $\dot{V}O_2$ on-kinetics and therefore increase performance in a subsequent trial could consequently not be verified. Finally to assess predicted versus actual 5-km run times, performance trials would have been beneficial for this work.

**Conclusions**

Due to dissimilar CS values and a high within-subject variability in *D´*, protocols using different exhaustive times within the currently recommended best-practice testing methodology cannot be used interchangeably. Compared to the shorter protocol (i.e. *Protocol II*), the longer protocol (i.e. *Protocol I*) resulted in significantly lower estimates of CS in a cohort of well-trained athletes. We therefore recommend the

consistent use of a particular testing protocol (i.e. iso-duration $t_{lim}$). Furthermore, to ensure same combined $t_{lim}$ throughout all tests, we recommended the use of fixed-durations over fixed-distances.

Acknowledgements: The corresponding author wants to express his thank to the Austrian Institute of Sports Medicine for their courtesy using their laboratory equipment.

References:

Bergstrom, H. C., Housh, T. J., Cochrane-Snyman, K. C., Jenkins, N. D., Byrd, T., Switalla, J. R., . . . Johnson, G. O. (2017). A Model for Identifying Intensity Zones above Critical Velocity. *Journal of Strength and Conditioning Research*. doi: 10.1519/JSC.0000000000001769

Bishop, D., Jenkins, D. G., & Howard, A. (1998). The critical power function is dependent on the duration of the predictive exercise tests chosen. *International Journal of Sports Medicine, 19*(2), 125-129. doi: 10.1055/s-2007-971894

Bland, J. M., & Altman, D. G. (1986). Statistical methods for assessing agreement between two methods of clinical measurement. *Lancet, 1*(8476), 307-310

Broxterman, R. M., Ade, C. J., Poole, D. C., Harms, C. A., & Barstow, T. J. (2013). A single test for the determination of parameters of the speed-time relationship for running. *Respiratory Physiology and Neurobiology, 185*(2), 380-385. doi: 10.1016/j.resp.2012.08.024

Buchheit, M. (2016). The Numbers Will Love You Back in Return-I Promise. *International Journal of Sports Physiology and Performance, 11*(4), 551-554. doi: 10.1123/IJSPP.2016-0214

Bull, A. J., Housh, T. J., Johnson, G. O., & Rana, S. R. (2008). Physiological responses at five estimates of critical velocity. *European Journal of Applied Physiology, 102*(6), 711-720. doi: 10.1007/s00421-007-0649-7

Burnley, M., Davison, G., & Baker, J. R. (2011). Effects of priming exercise on VO2 kinetics and the power-duration relationship. *Medicine and Science in Sports and Exercise, 43*(11), 2171-2179. doi: 10.1249/MSS.0b013e31821ff26d

Burnley, M., Doust, J. H., & Jones, A. M. (2005). Effects of prior warm-up regime on severe-intensity cycling performance. *Medicine and Science in Sports and Exercise, 37*(5), 838-845

Busso, T., Gimenez, P., & Chatagnon, M. (2010). A comparison of modelling procedures used to estimate the power-exhaustion time relationship. *European Journal of Applied Physiology, 108*(2), 257-263. doi: 10.1007/s00421-009-1213-4

Dekerle, J., de Souza, K. M., de Lucas, R. D., Guglielmo, L. G., Greco, C. C., & Denadai, B. S. (2015). 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*(9), e0138428. doi: 10.1371/journal.pone.0138428

di Prampero, P. E. (1999). The concept of critical velocity: a brief analysis. *European Journal of Applied Physiology and Occupational Physiology, 80*(2), 162-164. doi: 10.1007/s004210050574

Faul, F., Erdfelder, E., Buchner, A., & Lang, A. G. (2009). Statistical power analyses using G*Power 3.1: tests for correlation and regression analyses. *Behaviour Research Methods, 41*(4), 1149-1160. doi: 10.3758/BRM.41.4.1149

Ferguson, C., Rossiter, H. B., Whipp, B. J., Cathcart, A. J., Murgatroyd, S. R., & Ward, S. A. (2010). Effect of recovery duration from prior exhaustive exercise on the

parameters of the power-duration relationship. *Journal of Applied Physiology, 108*(4), 866-874. doi: 10.1152/japplphysiol.91425.2008

Ferguson, C., Wilson, J., Birch, K. M., & Kemi, O. J. (2013). Application of the speed-duration relationship to normalize the intensity of high-intensity interval training. *PLOS ONE, 8*(11), e76420. doi: 10.1371/journal.pone.0076420

Florence, S., & Weir, J. P. (1997). Relationship of critical velocity to marathon running performance. *European Journal of Applied Physiology and Occupational Physiology, 75*(3), 274-278

Galbraith, A., Hopker, J., Lelliott, S., Diddams, L., & Passfield, L. (2014). A single-visit field test of critical speed. *International Journal of Sports Physiology and Performance, 9*(6), 931-935. doi: 10.1123/ijspp.2013-0507

Galbraith, A., Hopker, J., & Passfield, L. (2015). Modeling Intermittent Running from a Single-visit Field Test. *International Journal of Sports Medicine, 36*(5), 365-370. doi: 10.1055/s-0034-1394465

Galbraith, A., Hopker, J. G., Jobson, S. A., & Passfield, L. (2011). A novel field test to determine critical speed. *Journal of Sports Medicine & Doping Studies, 01*(01), 1-4

Hill, D. W. (1993). The critical power concept. A review. *Sports Medicine, 16*(4), 237-254

Hill, D. W., Vingren, J. L., Nakamura, F. Y., & Kokobun, E. (2011). Relationship between speed and time in running. *International Journal of Sports Medicine, 32*(7), 519-522. doi: 10.1055/s-0031-1275298

Hinckson, E. A., & Hopkins, W. G. (2005). Reliability of time to exhaustion analyzed with critical-power and log-log modeling. *Medicine and Science in Sports and Exercise, 37*(4), 696-701

Hopkins, W. G., & Hewson, D. J. (2001). Variability of competitive performance of distance runners. *Medicine and Science in Sports and Exercise, 33*(9), 1588-1592

Housh, T. J., Devries, H. A., Housh, D. J., Tichy, M. W., Smyth, K. D., & Tichy, A. M. (1991). The relationship between critical power and the onset of blood lactate accumulation. *Journal of Sports Medicine and Physical Fitness, 31*(1), 31-36

Hughson, R. L., Orok, C. J., & Staudt, L. E. (1984). A high velocity treadmill running test to assess endurance running potential. *International Journal of Sports Medicine, 5*(1), 23-25. doi: 10.1055/s-2008-1025875

Jenkins, D. G., Kretek, K., & Bishop, D. (1998). The duration of predicting trials influences time to fatigue at critical power. *Journal of Science and Medicine in Sport, 1*(4), 213-218

Jenkins, D. G., & Quigley, B. M. (1992). Endurance training enhances critical power. *Medicine and Science in Sports and Exercise, 24*(11), 1283-1289

Jones, A. M., Vanhatalo, A., Burnley, M., Morton, R. H., & Poole, D. C. (2010). Critical power: implications for determination of V O2max and exercise tolerance. *Medicine and Science in Sports and Exercise, 42*(10), 1876-1890. doi: 10.1249/MSS.0b013e3181d9cf7f

Karsten, B., Baker, J., Naclerio, F., Klose, A., Bianco, A., & Nimmerichter, A. (2017). Time Trials versus Time to Exhaustion Tests: Effects on Critical Power, W' and Oxygen Uptake Kinetics. *International Journal of Sports Physiology and Performance*, 1-22. doi: 10.1123/ijspp.2016-0761

Karsten, B., Hopker, J., Jobson, S. A., Baker, J., Petrigna, L., Klose, A., & Beedie, C. (2017). Comparison of inter-trial recovery times for the determination of critical

power and W' in cycling. *Journal of Sports Sciences*, *35*(14), 1420-1425. doi: 10.1080/02640414.2016.1215500

Karsten, B., Jobson, S. A., Hopker, J., Jimenez, A., & Beedie, C. (2014). High agreement between laboratory and field estimates of critical power in cycling. *International Journal of Sports Medicine, 35*(4), 298-303. doi: 10.1055/s-0033-1349844

Karsten, B., Jobson, S. A., Hopker, J., Stevens, L., & Beedie, C. (2015). Validity and reliability of critical power field testing. *European Journal of Applied Physiology*, *115*(1), 197-204. doi: 10.1007/s00421-014-3001-z

Karsten, B., Larumbe-Zabala, E., Kandemir, G., Hazir, T., Klose, A., & Naclerio, F. (2016). The Effects of a 6-Week Strength Training on Critical Velocity, Anaerobic Running Distance, 30-M Sprint and Yo-Yo Intermittent Running Test Performances in Male Soccer Players. *PLOS ONE*, *11*(3), e0151448. doi: 10.1371/journal.pone.0151448

Kranenburg, K. J., & Smith, D. J. (1996). Comparison of critical speed determined from track running and treadmill tests in elite runners. *Medicine and Science in Sports and Exercise*, *28*(5), 614-618

Nimmerichter, A., Novak, N., Triska, C., Prinz, B., & Breese, B. C. (2017). Validity of Treadmill-Derived Critical Speed on Predicting 5000-Meter Track-Running Performance. *Journal of Strength and Conditioning Research, 31*(3), 706-714. doi: 10.1519/JSC.0000000000001529

Nimmerichter, A., Steindl, M., & Williams, C. A. (2015). Reliability of the Single-Visit Field Test of Critical Speed in Trained and Untrained Adolescents. *Sports*, *3*(4), 358-368. doi: 10.3390/sports3040358

Poole, D. C., Ward, S. A., Gardner, G. W., & Whipp, B. J. (1988). Metabolic and respiratory profile of the upper limit for prolonged exercise in man. *Ergonomics, 31*(9), 1265-1279. doi: 10.1080/00140138808966766

Sawyer, B. J., Morton, R. H., Womack, C. J., & Gaesser, G. A. (2012). V O2max may not be reached during exercise to exhaustion above critical power. *Medicine and Science in Sports and Exercise, 44*(8), 1533-1538. doi: 10.1249/MSS.0b013e31824d2587

Triska, C., Karsten, B., Nimmerichter, A., & Tschan, H. (2017). Iso-duration Determination of D' and CS under Laboratory and Field Conditions. *International Journal of Sports Medicine, 38*(7), 527-533. doi: 10.1055/s-0043-102943

Triska, C., Tschan, H., Tazreiter, G., & Nimmerichter, A. (2015). Critical Power in Laboratory and Field Conditions Using Single-visit Maximal Effort Trials. *International Journal of Sports Medicine, 36*(13), 1063-1068. doi: 10.1055/s-0035-1549958

Van Cutsem, J., Marcora, S., De Pauw, K., Bailey, S., Meeusen, R., & Roelands, B. (2017). The Effects of Mental Fatigue on Physical Performance: A Systematic Review. *Sports Medicine, 47*(8), 1569-1588. doi: 10.1007/s40279-016-0672-0

Vandewalle, H., Vautier, J. F., Kachouri, M., Lechevalier, J. M., & Monod, H. (1997). Work-exhaustion time relationships and the critical power concept. A critical review. *Journal of Sports Medicine and Physical Fitness, 37*(2), 89-102

Table 1. Estimates of CS and *D´* (mean ± SD) resulting from *Protocol I* and *Protocol II*.

Table 2. Resting and post-exercise blood lactate concentrations ($mmol \cdot L^{-1}$) and $HR_{max}$ post-exercise ($beats \cdot min^{-1}$) (mean ± SD).

Figure 1. Relationship of CS and *D´* between protocols (panels a and b). The dotted line represents the linear regression and the solid grey line represents the line of identity. Bland-Altman plots of the differences between the protocols of CS and *D´* (panels c and d).

Figure 2. Speed-duration relationship for *Protocol I* (black) and *Protocol II* (grey). The data points represent the mean speeds and the error bars the standard deviation.