ORIGINAL ARTICLE

Anaerobic work capacity in cycling: the effect of computational method

Erik P. Andersson[1,2] · Philipp Bachl[1,3] · Anna Schmuttermair[3] · Craig A. Staunton[1] · Thomas L. Stöggl[3,4]

Abstract

Purpose To compare the anaerobic work capacity (AnWC, i.e., attributable anaerobic mechanical work) assessed using four different approaches/models applied to time-trial (TT) cycle-ergometry exercise.

Methods Fifteen male cyclists completed a 7×4-min submaximal protocol and a 3-min all-out TT (TT_{AO}). Linear relationships between power output (PO) and submaximal metabolic rate were constructed to estimate TT-specific gross efficiency (GE) and AnWC, using either a measured resting metabolic rate as a Y-intercept ($7 + Y_{LIN}$) or no measured Y-intercept ($7 - Y_{LIN}$). In addition, GE of the last submaximal bout (GE_{LAST}) was used to estimate AnWC, and critical power (CP) from TT_{AO} ($CP_{3'AO}$) was used to estimate mechanical work above CP (W', i.e., "AnWC").

Results Average PO during TT_{AO} was 5.43 ± 0.30 and CP was 4.48 ± 0.23 W•kg^{-1}. The TT-associated GE values were ~22.0% for both $7 + Y_{LIN}$ and $7 - Y_{LIN}$ and ~21.1% for GE_{LAST} (both $P < 0.001$). The AnWC were 269 ± 60, 272 ± 55, 299 ± 61, and 196 ± 52 J•kg^{-1} for the $7 + Y_{LIN}$, $7 - Y_{LIN}$, GE_{LAST}, and $CP_{3'AO}$ models, respectively ($7 + Y_{LIN}$ and $7 - Y_{LIN}$ versus GE_{LAST}, both $P < 0.001$; $7 + Y_{LIN}$, $7 - Y_{LIN}$, and GE_{LAST} versus $CP_{3'AO}$, all $P < 0.01$). For the three pair-wise comparisons between $7 + Y_{LIN}$, $7 - Y_{LIN}$, and GE_{LAST}, typical errors in AnWC values ranged from 7 to 11 J•kg^{-1}, whereas $7 + Y_{LIN}$, $7 - Y_{LIN}$, and GE_{LAST} versus $CP_{3'AO}$ revealed typical errors of 55–59 J•kg^{-1}.

Conclusion These findings demonstrate a substantial disagreement in AnWC between $CP_{3'AO}$ and the other models. The $7 + Y_{LIN}$ and $7 - Y_{LIN}$ generated 10% lower AnWC values than the GE_{LAST} model, whereas $7 + Y_{LIN}$ and $7 - Y_{LIN}$ generated similar values of AnWC.

Keywords All-out pacing · Maximal accumulated oxygen deficit method · Metabolic demand · Time trial · Reliability · Supramaximal exercise

✉ Erik P. Andersson
  erik.andersson@miun.se

1 Swedish Winter Sports Research Centre, Department of Health Sciences, Mid Sweden University, Östersund, Sweden

2 School of Sport Sciences, Faculty of Health Sciences, UiT The Arctic University of Norway, Tromsø, Norway

3 Department of Sport and Exercise Science, University of Salzburg, Salzburg, Austria

4 Red Bull Athlete Performance Center, Salzburg, Austria

Abbreviations

$7 + Y_{LIN}$	The linear regression model between power output and metabolic rate used to estimate the required instantaneous metabolic rate during the time trial, including a baseline value of metabolic rate as a Y-intercept
$7 - Y_{LIN}$	The linear regression model between power output and metabolic rate used to estimate the required instantaneous metabolic rate during the time trial, excluding a baseline value of metabolic rate as a Y-intercept
AnC	Anaerobic capacity
AnWC	Anaerobic work capacity
MR_{BL}	Baseline metabolic rate
CO_2	Carbon dioxide
CP	Critical power

$CP_{3'AO}$	The critical power method for determining anaerobic work capacity based on the 30-s average end power during the 3-min all-out time trial as critical power
ES_{Hg_av}	Hedges' g effect size
GE	Gross efficiency
GE_{LAST}	Gross efficiency from the last submaximal stage used to estimate the required instantaneous metabolic rate during the time trial
GE_{REG}	Gross efficiency calculated from the linear regression equation between power output and metabolic rate
MAOD	Maximal accumulated oxygen deficit
MR	Metabolic rate
MR_{AE}	Aerobic metabolic rate
MR_{AN}	Anaerobic metabolic rate
MR_{TT_req}	The required metabolic rate during the time trial
O_2	Oxygen
PO	Power output
PO_{AN_cont}	Anaerobic contribution to power output
PO_{TT}	Power output during the time trial
RER	Respiratory exchange ratio
TT	Time trial
TT_{AO}	The 3-min all-out time trial
$\dot{v}o_2$	Oxygen consumption
$\dot{v}o_{2max}$	Maximal oxygen consumption
$\dot{v}o_{2peak}$	Peak oxygen consumption
W′	Mechanical work above critical power

Introduction

Currently, there is no gold standard method for determining anaerobic capacity (AnC) or anaerobic work capacity (AnWC; i.e., the amount of anaerobic energy that is converted to external work) during whole-body exercise (Noordhof et al. 2010). The most commonly used method for estimating AnC is the linear regression method also referred to as the maximal accumulated oxygen deficit (MAOD) method (Medbø et al. 1988). This method is based on the assumption of a linear relationship between external exercise intensity (e.g., speed or power output) and oxygen uptake ($\dot{v}o_2$) during submaximal steady-state exercise. Subsequently, the linear relationship can be extrapolated to predict the metabolic requirement in $\dot{v}o_2$ equivalents at intensities above maximal $\dot{v}o_2$ ($\dot{v}o_{2max}$), with the difference between the required accumulated VO_2, and the measured accumulated VO_2, representing the oxygen deficit. However, due to the different energy equivalents for fat and carbohydrate oxidation, and the effect of submaximal exercise intensity on substrate utilization (Shaw et al. 2014), a linear regression method based on external power output (PO) and metabolic

rate (MR) (linear PO-MR regression method) has been suggested to be more appropriate than the traditional MAOD method for determining the required total MR and AnC during supramaximal exercise (Andersson and McGawley 2018; Andersson et al. 2020). Such a method would also enable instantaneous calculation of gross efficiency (GE) during supramaximal exercise through the use of the regression equation to determine the required total MR (Andersson et al. 2020).

Another common method used to determine AnWC and/or AnC is the GE method (Andersson et al. 2020; Noordhof et al. 2011; Serresse et al. 1988). This method requires a single stage of exercise at submaximal steady-state intensity, just below the second ventilatory threshold, as well as a supramaximal exercise bout. For the conventional GE method, the anaerobic contribution to PO is determined as the difference between PO and the aerobic contribution to PO (calculated as aerobic MR multiplied by GE) with AnWC calculated as the anaerobic contribution to PO integrated over time (Noordhof et al. 2011). Thus, AnWC is dependent on both the AnC and GE. Although a not common practice, the GE method can easily determine AnC. This is performed by dividing supramaximal PO by GE to calculate the required total MR, with anaerobic MR being calculated as the difference between required total MR and aerobic MR, which when integrated over time represents the AnC (Andersson et al. 2020).

The obvious advantage of the GE versus the linear PO-MR regression method is that it is far less time-consuming, as it only requires one submaximal stage compared to the linear regression method which requires ~5–10 submaximal stages (Noordhof et al. 2011). However, one assumption with the GE method that differs from the linear PO-MR regression method is that GE is PO independent, which only is the case for the linear PO-MR regression method if the Y-intercept value in the regression is zero (Andersson et al. 2020). In addition, the GE method can be converted to a linear PO-MR equation where the slope represents the reciprocal value of GE combined with a zero Y-intercept (Andersson et al. 2020). Although the traditional MAOD method is inappropriate for determining AnWC (i.e., anaerobically attributable mechanical work) (Medbø et al. 1988), the supramaximal instantaneous GE calculated from the linear PO-MR regression method can be used to determine the anaerobically attributable PO and, thus, AnWC during supramaximal exercise (Andersson et al. 2020). One clear difference between the GE method and the linear PO-MR regression method is that the GE method is solely based on a constant GE. In contrast, the linear PO-MR regression method assumes an increasing GE if the Y-intercept of the linear PO-MR regression is positive, which is the case for cycle-ergometry exercise (Ettema and Lorås 2009). Therefore, the associated GE during supramaximal cycle exercise

would likely be higher for the linear PO-MR regression method compared to the GE method and result in higher values of estimated AnWC/AnC for the latter method.

Another concept that can be used to differentiate between aerobic and anaerobic power contributions to external PO is the critical power (CP) method. With the CP concept, the CP threshold equals the maximal PO that is generated by primarily aerobic energy sources and can, at least in theory, be maintained indefinitely, whereas the exercise duration above CP is finite (Vanhatalo et al. 2007, 2011). For determining CP, the hyperbolic relationship between exercise duration and PO for maximal exercise first needs to be established. This requires participants to complete ≥ 4 separate maximal tests to exhaustion at several different fixed power outputs (Vanhatalo et al. 2011). In contrast to this traditional and time-consuming approach of determining CP, a modified version of the CP protocol has been proposed, which involves just a single 3-min all-out time trial (TT) (Vanhatalo et al. 2007). Due to depletion of the anaerobic energy reserve during the initial stages of exercise (≤ 2 min of exercise), the average PO during the final 30 s (referred to as "end power") of the 3-min all-out test is considered to represent CP (Vanhatalo et al. 2007). With the CP method, the total mechanical work above critical power (W') is calculated as PO above CP integrated over time and is often referred to as a marker of AnWC (Hill 1993; Morton 2006). However, referring to W' as a pure AnWC may be incorrect as it also can include aerobically attributable work (Vinetti et al., 2017, 2019). The aerobic component of W' is, though, compensated by the assumption that PO at CP is exclusively supplied by aerobic metabolism at the onset of exercise (i.e., neglecting the primary component of $\dot{v}o_2$ kinetics) with the implication that W' can represent a valid AnWC.

In a previous study by Hill and Smith (1993), the W' estimate of AnWC was compared with an adopted MAOD method, with similar average estimates of AnWC found between the two methods. However, these estimates of AnWC were not perfectly correlated ($r = 0.55$ and $r = 0.78$ for females and males, respectively). Although, Hill and Smith (1993) found comparable average values of AnWC for the CP and GE methods, a more recent study by Dekerle et al. (2006) showed a poor agreement between W' based on CP and AnWC during a 90-s all-out cycling test, with AnWC being calculated as the difference between external PO and aerobically attributable PO integrated over time. Several possible factors could explain the disagreement in AnWC determined with the CP (i.e., W') and "MAOD"/"GE" methods used in previous studies (Dekerle et al. 2006; Hill and Smith 1993). One main factor for these divergences is likely related to the fact that aerobically attributable PO is approximated with the CP model without any physiological assessment. However, the W' measure of AnWC based on the CP concept can attain a lower, similar, or even higher, value

than the AnWC based on the GE method because there can be aerobically attributable PO above CP up to the maximal aerobically attributable PO (i.e., PO at $\dot{v}o_{2\,max}$) (Vinetti et al. 2017, 2019). For instance, if the average fractional utilization of $\dot{v}o_{2\,max}$ (as a percentage of $\dot{v}o_{2max}$) during a maximal effort is the same as the corresponding $\dot{v}o_2$ requirement at CP (as a percentage of $\dot{v}o_{2max}$), W' would, in theory, attain a similar value as AnWC based on the GE method.

Only a few studies have compared the GE and MAOD computational methods for determining AnC during cycle ergometry (Noordhof et al. 2011) and treadmill roller-skiing exercise (Andersson and McGawley 2018; Andersson et al. 2020). In each of these studies, the disagreement between AnC estimates across different models was found to be considerable and suggests that different models should not be used interchangeably (Andersson et al. 2020; Noordhof et al. 2011). However, there is sparse research that has compared the W' estimate of "AnWC" (based on the CP method) against the AnWC estimated with the MAOD and/or GE methods (Dekerle et al. 2006; Hill and Smith 1993). To our knowledge, there is only one study that has analyzed the agreement between W' based on CP against the estimated accumulated oxygen deficit based on a linear $\dot{v}o_2$-PO regression method during a 3-min all-out time trial (Muniz-Pumares et al. 2016). One limitation of this previous study was that the agreement between the two methods was assessed by correlational analysis (Bland and Altman 1999) and that different units were correlated, i.e., correlating AnC (in oxygen equivalents) with AnWC (Winter and Fowler 2009).

Given the limited number of methodological studies, the primary aim of the current study was to compare estimates of AnWC and/or AnC generated during a 3-min all-out time trial (TT_{AO}) using four different models: the 7×4-min linear regression method based on PO and MR with, and without, the inclusion of a resting baseline MR value as a Y-intercept ($7 + Y_{LIN}$ and 7-Y_{LIN}, respectively); the GE method using the last exercise intensity (GE_{LAST}); and the CP method based on the 30-s average end power as CP ($CP_{3'AO}$).

The hypotheses of the current study were as follows:

1) the GE_{LAST} model would generate the highest values of AnWC/AnC compared to the other models tested; and 2) there would be disagreement in W'/AnWC between the $CP_{3'AO}$ versus the $7 + Y_{LIN}$, 7-Y_{LIN}, and GE_{LAST} models.

Methods

Participants

Fifteen highly trained male cyclists (mean $\pm$ standard deviation: age: 28.0 ± 4.7 yr., body mass: 78.5 ± 7.7 kg, stature:

183.5 ± 7.0 cm) were recruited for this study (Tier 3–4, McKay et al. 2022). During the 24 h prior to testing, participants were instructed to perform ≤ 2 h of low-intensity exercise and refrain from intake of alcohol. The last regular meal had to be eaten at least 3 h pre-test; however, a small snack (e.g., an energy bar) could be consumed 1–2 h before the test. The participants could consume water ad libitum, but no intake of any food was allowed during the test. The study was approved by the ethical review board of the University of Salzburg (EK-GZ: 05/2020). All participants were fully informed about the nature of the study and provided written consent before the first test. Exclusion criteria were any of the following: $\dot{v}o_{2peak} < 55$ ml kg^{-1} min^{-1}; no previous experience of laboratory cycle-ergometry tests; and injury or illness.

Equipment, measurements, and testing procedures

The participants used their own clothes, cycling shoes, cleats, and pedals. All testing was performed on a bike designed for TT tests (Monark LC7TT, Monark Exercise AB, Vansbro, Sweden) equipped with road race handlebars and standard shifting mechanics (Shimano Ultegra 11 Speed, Shimano Inc., Osaka, Japan). Cycling PO was logged continuously as second-by-second data. The bike was fully adjustable to the rider's preferences (i.e., seat tube length, saddle-setback, the height of the stem, and forward length of the stem). The participant's stature and body mass were measured before the first test using an electronic scale (Seca 764, Hamburg, Germany). Respiratory variables were measured using a Cosmed Quark CPET mixing chamber ergospirometry system (Cosmed Srl, Rome, Italy) with raw data as 10-s values. This set-up was used to provide valid and reliable metabolic measurements, especially at high ventilation rates in highly trained athletes (Nieman et al.

2013; Winkert et al. 2020). Prior to each test, the oxygen (O_2) and carbon dioxide (CO_2) sensors were calibrated using a two-point calibration procedure with ambient air conditions and the anticipated expiratory gas percentages using a known calibration gas containing 15% O_2 and 5% CO_2 (UN 1950 Aerosols, Cortex Biophysik GmbH, Leipzig, Germany). The flow volume was calibrated using a 3-L syringe (M9424, Medikro Oy, Kuopio, Finland). Blood lactate concentration was measured via ear-lobe capillary samples (20 µL) that were subsequently analyzed using a Biosen S-line (EKF diagnostic GmbH, Magdeburg, Germany). The Biosen S-line was calibrated with a known standard solution of 12 mmol L^{-1}. The blood samples were collected at the final minute of the 6-min warm-up, directly after completion of the submaximal protocol, 1 min prior to the start of the TT, and 2 min after the TT (see Fig. 1). Heart rate was measured continuously throughout the test using a heart rate belt (Wahoo Kickr, Wahoo Fitness, Atlanta, GA, United States) that was connected to the ergospirometry system.

A 3-min baseline respiratory sample was measured prior to the warm-up with the participant being seated upright on the bicycle with both feet resting on the pedals. The baseline respiratory sample was preceded by a 3-min seated rest. The warm-up protocol consisted of 6-min at ~40% of peak $\dot{v}o_2$ ($\dot{v}o_{2peak}$) and started immediately following the baseline respiratory sample. The warm-up was directly followed by the submaximal protocol that consisted of seven 4-min stages (7 × 4) that were performed between ~40–74% of $\dot{v}o_{2peak}$ with ~6% increment per stage. This was followed by a 6-min passive rest and the 3-min TT_{AO}. The TT_{AO} concept was utilized to establish CP, using the 3-min all-out method, which allows for an estimation of AnWC (Burnley et al. 2006; Vanhatalo et al. 2007). During the TT_{AO}, heart rate, PO, and elapsed time were not visible to the participant and the average PO of the last 30 s of the all-out effort was used

Fig. 1 A schematic overview of the testing protocol where participants performed a 3-min all-out time trial (TT_{AO}). After a 3-min baseline measure of oxygen consumption (3′ B) and a 6-min warm-up (6′ W-up), the 7×4-min submaximal exercise stages were performed and followed by a 6-min passive break (6′ PB). Capillary blood samples for the determination of blood lactate concentration (La⁻) were collected prior to and immediately after the submaximal stages and 1-min before and 2-min after the TT. Abbreviations: ′, minute; $\dot{v}o_{2peak}$, peak oxygen consumption; SUB, sub-maximal
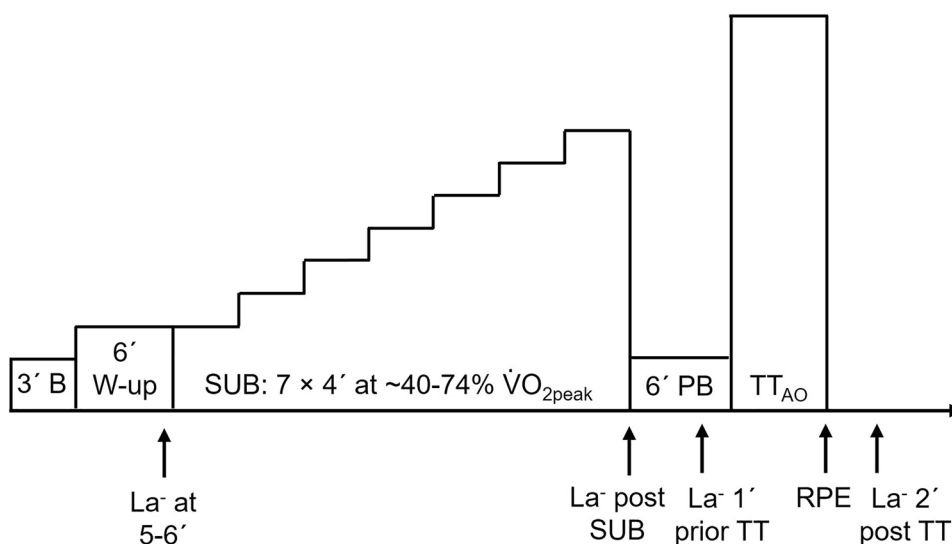
to calculate CP according to previous studies (Burnley et al. 2006; Vanhatalo et al. 2007). Furthermore, participants were instructed to provide maximal effort ("no holding back") from the beginning of the TT_{AO}. The PO for the warm-up and submaximal stages was fixed and cadence independent whereas, for the TT_{AO}, the PO was regulated freely by the athlete via the bike's shifters and was cadence dependent. The cadence was self-selected during the submaximal stages and TT_{AO}. The cadence for the submaximal exercise was determined as the average of the final minute of each stage. Peak cadence and peak PO during TT_{AO} were determined as the highest respective 5-s moving average. A schematic overview of the testing procedure is presented in Fig. 1. During the final minute of the last submaximal stage and immediately after the TT, the participant reported a rating of perceived exertion using the Borg 6–20 scale (Borg 1982). All participants were familiar with submaximal and maximal cycle ergometry efforts. The submaximal power outputs were selected based on previous test results.

Processing of respiratory data

Respiratory and heart rate data were collected continuously during the submaximal exercise and the TT_{AO}. To enable a higher resolution of respiratory data during TT_{AO} (i.e., to obtain a more dynamic respiratory response), raw 10-s respiratory data were transferred to second-by-second data (using piecewise constant interpolation for each 10-s value) and smoothed using a 9-s counterbalanced moving average (i.e., ± 4-s time window for smoothing), which was conducted twice. For the start-point of the TT_{AO}, a gradual increase in the smoothing function time window was used up to the fifth second, whereafter the 9-s counterbalanced moving average was used. For the endpoint of TT_{AO}, the same principle was employed, but with a gradual decrease in the smoothing time window over the last four seconds. The highest 20-s moving average during the TT_{AO} was used to calculate $\dot{v}o_{2peak}$, with peak RER taken over the same period. In addition, $\dot{v}o_{2peak}$ was converted to a peak aerobic metabolic rate [Eq. 1] using an RER of 1.00 (i.e., assuming 100% carbohydrate utilization at $\dot{v}o_{2peak}$). Peak heart rate was considered the highest 10-s average value.

Calculations

Submaximal exercise

Energy expenditure was calculated from $\dot{v}o_2$ and RER ($\dot{v}co_2 \dot{v}o_2^{-1}$) according to the equation published by Weir (1949) and then converted into a MR. The MR was based on the average $\dot{v}o_2$ in L min^{-1} and RER values (≤ 1.00) during the final minute of each stage of the submaximal exercise protocol.

$$MR[W] = \frac{4184(\dot{V}O_2(1.1RER + 3.9))}{60} \tag{1}$$

GE was calculated as:

$$GE = \frac{PO[W]}{MR[W]} \tag{2}$$

Net efficiency was calculated as:

$$Net\ efficiency = \frac{PO[W]}{MR - MR_{BL}[W]} \tag{3}$$

where MR_{BL} is the baseline MR calculated from an average 3-min baseline $\dot{v}o_2$ and RER measurement with the participant seated on the cycle ergometer with no pedaling. Delta efficiency was calculated as the reciprocal value of the slope of the PO-MR regression equation. Neither net efficiency nor delta efficiency was used for estimating AnWC/AnC.

Estimating AnC, AnWC, and supramaximal GE

A linear relationship between PO (W) and MR (W) during the final minute of each of the 7×4-min submaximal stages was derived for each participant with the baseline MR (i.e., the MR at zero speed) included ($7 + Y_{LIN}$) or excluded ($7-Y_{LIN}$) from the model. In the latter case, the Y-intercept was based on all data points in the regression but excluding the baseline resting value of MR. The $7 + Y_{LIN}$ and $7-Y_{LIN}$ regression equations were used to estimate the required instantaneous MR during the 3-min TT (MR_{TT_req}) at each 1-s time point. Submaximal GE from the last submaximal stage (GE_{LAST}) was also used to estimate the MR_{TT_req} at each 1-s time point of the TT. MR_{TT_req} was calculated by dividing instantaneous PO with GE_{LAST}. For illustrative purposes, GE_{LAST} was converted to a PO-MR linear equation, where the slope was calculated as the reciprocal value of GE_{LAST}, with a Y-intercept of zero due to the constant GE assumption for the GE_{LAST} model. The instantaneous second-by-second GE values during the TT were also calculated for the $7 + Y_{LIN}$ and $7-Y_{LIN}$ models (GE_{REG}) as instantaneous PO divided by the instantaneous required MR calculated from the regression equation.

For the $7 + Y_{LIN}$, $7-Y_{LIN}$, and GE_{LAST} models, the instantaneous anaerobic MR (MR_{AN}) at each 1-s time-point (t) of the TT could then be expressed as:

$$MR_{AN,t}[W] = MR_{TT\_req,t} - MR_{AE,t} \tag{4}$$

where MR_{AE} is the aerobic MR calculated according to Eq. 1 and using an RER of 1.00 (i.e., assuming 100% carbohydrate utilization during the TT).

For all three models, the total anaerobic energy production (i.e., AnC [kJ kg^{-1}]) was calculated by integrating MR$_{an}$ over the 3-min TT.

Anaerobic PO contribution (PO$_{AN\_cont}$) (i.e., PO attributable to MR$_{AN}$) at each 1-s time point (t) of the TT was calculated for the 7 + Y$_{LIN}$ and 7-Y$_{LIN}$ models as:

$$PO_{AN_{cont},t} = PO_{TT,t}[W] - (MR_{AE,t} \times GE_{REG,t}) \quad (5)$$

where PO$_{TT}$ is the PO during the TT.

For GE$_{LAST}$, the same equation was used but with the exception that GE$_{REG}$ was changed to GE$_{LAST}$.

The AnWC in joules (J) was calculated for the 7 + Y$_{LIN}$, 7-Y$_{LIN}$, and GE$_{LAST}$ models by integrating the model-specific PO$_{AN\_cont}$ (W) over the TT duration (s).

For the CP$_{3'AO}$ model, the total mechanical work above CP (i.e., W') was calculated as the PO above CP integrated over time. The CP was determined as the average PO of the last 30 s of TT$_{AO}$ (Vanhatalo et al. 2007). In the current study, W' based on CP$_{3'AO}$ has been referred to as an "AnWC", because W' is usually referred to as a surrogate marker of AnWC (Hill, 1993; Morton, 2006).

To calculate the average supramaximal GE during the TT for both the 7 + Y$_{LIN}$ and 7-Y$_{LIN}$ models, the estimated instantaneous GE of the 3-min TT was calculated as instantaneous PO divided by instantaneous MR$_{TT\_req}$ (derived from the linear PO-MR regression equation) and expressed as an average TT value.

Statistical analyses

All statistical tests were processed using Office Excel 2016 (Microsoft Corporation, Redmond, WA, USA) and the Statistical package for the social sciences (SPSS 25, IBM Corp., Armonk, NY, USA). The level of statistical significance was set at $\alpha = 0.05$. Normality of data was confirmed by visual inspection of Q–Q plots and histograms together with the Shapiro–Wilk analysis. Accordingly, data are presented as mean ± standard deviation, except in the case of peak heart rate and rating of perceived exertion where data are presented as median and interquartile range. In addition, the different AnWC estimates are presented as mean and 95% confidence interval. The PO-MR relationships for the 7 + Y$_{LIN}$ and 7-Y$_{LIN}$ models were assessed using linear regression analyses. One-way repeated measures ANOVA tests were used to compare GE and net efficiency between the seven submaximal stages as well as the GE, required MR, and AnC associated with the TT. A paired t test was used to compare the linear regression coefficients for the 7 + Y$_{LIN}$ and 7-Y$_{LIN}$ models. The precision of the two linear regression equations was assessed with the standard error of the estimate. For the ANOVA tests, the assumption of sphericity was assessed using Mauchly's test. For violated sphericity, a Greenhouse–Geisser correction of the degrees of freedom was used (epsilon ≤ 0.75). Eta-squared effect size was reported for the one-way repeated measures ANOVA tests. Bonferroni α corrections were applied to all ANOVA tests.

The mean difference ± 95% limits of agreement for the comparison of the AnWC estimates were evaluated using Bland–Altman calculations (Bland and Altman 1999). The mean difference was tested with a paired-sample t-test and the standardized mean difference (Hedges' g_{av}, effect size [ES$_{Hg\_av}$]) was reported according to the equations presented by Lakens (2013). In addition, the methodological error was evaluated via the overall standard error of measurement calculated as the square root of the within-groups mean square error term in the repeated measures ANOVA and the absolute typical error for the separate pair-wise comparisons. The typical error was also expressed as a percentage, i.e., as a percentage of the grand mean.

Results

Physiological and cadence responses to submaximal exercise

The PO, cadence, physiological responses, and two types of efficiency (i.e., GE and net efficiency) at the seven submaximal stages are shown in Table 1. The blood lactate concentrations at the fifth minute of the warm-up, and immediately after the submaximal exercise, were 1.0 ± 0.2 and 2.7 ± 0.5 mmol L^{-1}, respectively. The rating of perceived exertion value at the last minute of the final submaximal stage was 15 (interquartile range = 15–16). The GE increased from the first to the last submaximal stage prior to the TT ($F_{2,32} = 42.8$, $P < 0.001$, eta-squared effect size = 0.754), whereas the net efficiency remained unchanged during all the submaximal stages prior to the TT ($F_{3,36} = 2.5$, $P = 0.085$, eta- squared effect size = 0.151).

Performance, cadence, and physiological responses during TT$_{AO}$

The average PO during TT$_{AO}$ was 5.43 ± 0.30 W kg^{-1} at an average cadence of 92 ± 9 rev min^{-1}. The peak PO was 9.3 ± 1.2 W kg^{-1} and the peak cadence was 127 ± 19 rev•min^{-1}. The average cadence during the last 30 s of the TT$_{AO}$ was 90 ± 14 rev min^{-1}. The v̇o$_{2peak}$ was 64 ± 6 ml kg^{-1} min^{-1} (5.0 ± 0.6 L min^{-1}) at an RER of 1.10 ± 0.10. The average v̇o$_2$ was 51 ± 6 ml kg^{-1} min^{-1}, which resulted in an average fractional utilization of $80 \pm 3\%$ of v̇o$_{2peak}$. The peak heart rate was 183 (interquartile range = 179–186) beats min^{-1}. The blood lactate concentrations measured 1 min prior to the TT and 2 min after the TT were 2.1 ± 0.5 and

Table 1 Mean ± standard deviation of power outputs, cadences, heart rates, cardiorespiratory variables, and efficiencies associated with the seven submaximal stages (SUB$_{1-7}$) of cycle ergometry exercise, as well as the seated resting baseline (BL$_{REST}$) data

	BL$_{REST}$	SUB$_1$	SUB$_2$	SUB$_3$	SUB$_4$	SUB$_5$	SUB$_6$	SUB$_7$
Power output (W kg^{-1})	0	1.62 ± 0.11	1.93 ± 0.14	2.24 ± 0.16	2.54 ± 0.18	2.85 ± 0.20	3.16 ± 0.23	3.47 ± 0.25
Cadence (rev min^{-1})	–	68 ± 6	70 ± 6	73 ± 6	76 ± 6	79 ± 6	82 ± 5	84 ± 6
Heart rate (% of TT$_{peak}$)	40 ± 5	61 ± 3	65 ± 3	70 ± 3	74 ± 3	79 ± 3	84 ± 2	89 ± 2
MR$_{AE}$ (W•kg^{-1})	2.0 ± 0.4	8.8 ± 0.8	9.8 ± 0.9	11.1 ± 1.0	12.4 ± 1.0	13.7 ± 1.1	15.0 ± 1.1	16.5 ± 1.4
MR$_{AE}$ (% of MR$_{AE\_peak}$)	9 ± 2	40 ± 4	44 ± 4	50 ± 5	56 ± 4	62 ± 5	68 ± 5	74 ± 5
$\dot{V}_E$ (L min^{-1})	16.7 ± 4.6	52.2 ± 6.9	56.0 ± 7.4	63.0 ± 8.4	71.7 ± 9.9	79.4 ± 10.7	89.2 ± 11.6	100.7 ± 14.8
$\dot{V}_E \dot{V}O_2^{-1}$	35.1 ± 4.7	25.5 ± 2.0	24.4 ± 1.8	24.4 ± 1.7	24.8 ± 2.1	25.0 ± 1.9	25.6 ± 1.9	26.4 ± 2.4
RER ($\dot{V}CO_2 • \dot{V}O_2^{-1}$)	0.85 ± 0.08	0.87 ± 0.04	0.86 ± 0.04	0.87 ± 0.04	0.88 ± 0.04	0.89 ± 0.04	0.90 ± 0.04	0.91 ± 0.04
Gross efficiency (%)	–	18.5 ± 1.5	19.7 ± 1.4	20.3 ± 1.4	20.5 ± 1.2	20.9 ± 1.2	21.0 ± 1.2	21.1 ± 1.1
Net efficiency (%)	–	24.0 ± 2.1	24.8 ± 2.1	24.7 ± 1.9	24.5 ± 1.5	24.6 ± 1.9	24.3 ± 1.7	24.0 ± 1.6

Abbreviations: TT_{peak} time trial peak value, MR_{AE} aerobic metabolic rate, MR_{AE_peak} peak aerobic metabolic rate during the time trial, $\dot{V}_E$ ventilation rate, $\dot{V}_E \cdot \dot{V}O_2^{-1}$ ventilatory equivalent for oxygen; RER, respiratory exchange ratio

11.5 ± 2.6 mmol L^{-1}. The median rating of perceived exertion value immediately after the TT was 20 (interquartile range = 19–20).

Comparison between the different models used to estimate AnWC/AnC

The mean PO-MR regression lines that were based on the seven submaximal stages for the 7 + Y$_{LIN}$ and 7-Y$_{LIN}$ models and calculated from GE based on the last submaximal stage (GE$_{LAST}$) with extrapolation up to the TT are displayed in Fig. 2A. The mean ± standard deviation values of directly measured GE and the GE calculated from the 7 + Y$_{LIN}$ and 7-Y$_{LIN}$ models are displayed in Fig. 2B. The mean instantaneous total PO, aerobic contribution to PO (based on GE$_{LAST}$), and CP are presented in Fig. 2C, with the total PO above CP integrated over time representing the *W'* (i.e., "AnWC") estimated with the CP$_{3'AO}$ model. The CP was 4.5 ± 0.2 W kg^{-1}, which was equivalent to 93 ± 10%, 93 ± 9%, and 96 ± 10% of the peak aerobic MR (i.e., the $\dot{v}o_{2peak}$ expressed as a MR) for the 7 + Y$_{LIN}$, 7-Y$_{LIN}$, and GE$_{LAST}$ models. The mean AnWC and 95% confidence interval together with individual data (colored symbols) are presented in Fig. 2D. It can be noted that the AnWC was considerably lower for the CP$_{3'AO}$ model compared to 7 + Y$_{LIN}$, 7-Y$_{LIN}$, and GE$_{LAST}$ models, while the GE$_{LAST}$ generated the highest values of AnWC.

The data presented in Table 2 show that the standard error of the estimate was larger for the 7 + Y$_{LIN}$ versus the 7-Y$_{LIN}$ regression model. The regression slope and Y-intercept values were similar for the 7 + Y$_{LIN}$ and 7-Y$_{LIN}$ models, which resulted in similar average GE values during the TT. The average GE during the TT was ~0.9 percentage points lower for the GE$_{LAST}$ than the 7 + Y$_{LIN}$ and 7-Y$_{LIN}$ models.

Individual PO-MR regression data as based on the 7 + Y$_{LIN}$ and 7-Y$_{LIN}$ models together with GE calculated from the two linear regression equations for the submaximal stages and the TT, and measured GE, as well as the estimated AnWC for the 7 + Y$_{LIN}$, 7-Y$_{LIN}$, and GE$_{LAST}$ models are shown in Fig. 3. Comparisons of the AnC/AnWC estimates from the 3-min TT$_{AO}$ using the different models are presented in Table 2 and Fig. 4. As shown in Fig. 4A, the 7 + Y$_{LIN}$ and 7-Y$_{LIN}$ generated similar mean values of AnWC and the typical errors between the models were relatively low. The GE$_{LAST}$ model generated significantly higher values (~12% higher) of AnWC than the 7 + Y$_{LIN}$ and 7-Y$_{LIN}$ models, due to the lower GE (see Table 2), and the typical errors were relatively low (2–3%) (Figs. 4B,C). The CP$_{3'AO}$ model generated significantly lower values of AnWC (~30% lower, on average) than the 7 + Y$_{LIN}$, 7-Y$_{LIN}$, and GE$_{LAST}$ models and the typical errors were high (23–26%) for all the respective comparisons (Figs. 4D–F).

The variation in Y-intercept values for the 7-Y$_{LIN}$ model was highly related to the variation in the AnWC estimates between the 7-Y$_{LIN}$ and GE$_{LAST}$ models (r^2 = 0.825; Fig. 5A), whereas the variation in Y-intercept values for the 7 + Y$_{LIN}$ model was non-significantly related to the variation in the AnWC estimates between the 7 + Y$_{LIN}$ and GE$_{LAST}$ models (r^2 = 0.052; Fig. 5B).

Discussion

The main findings were as follows: 1) the GE$_{LAST}$ model generated the highest AnWC/AnC estimate and the 7 + Y$_{LIN}$ and 7-Y$_{LIN}$ models generated similar average values of AnWC/AnC; 2) the typical errors in AnWC between the 7 + Y$_{LIN}$, 7-Y$_{LIN}$ and GE$_{LAST}$ models were low; 3) the AnWC estimated with the CP$_{3'AO}$ model was substantially lower and

Fig. 2 Data based on 7 stages of submaximal exercise followed by ▶ a 3-min supramaximal all-out time trial (TT_{AO}) (A-D). **A** The two regression models between mean ± standard deviation power output (PO) and metabolic rate during 7×4-min stages of continuous cycle ergometry exercise and the regression line calculated based on the gross efficiency (GE) from the last submaximal stage (GE_{LAST}) together with the estimated total metabolic requirements (diamonds) at the average PO attained during the 3-min time trial. The red line is the linear relationship when using a Y-intercept ($7+Y_{LIN}$) for baseline metabolic rate, the black solid line when excluding a Y-intercept value ($7\text{-}Y_{LIN}$), and the yellow line is the regression line based on GE_{LAST} (i.e., with the slope being the reciprocal value of GE); **B** GE as mean ± standard deviation for the seven 4-min stages of submaximal cycling and GE calculated from the two regression equations (i.e., $7+Y_{LIN}$ and $7\text{-}Y_{LIN}$) for the submaximal stages and the TT; (C) Total PO (PO) and aerobic power contribution (based on GE_{LAST}) presented as second-by-second average time-trial data and the estimated critical power (CP); (D) The mean anaerobic work capacity (AnWC) and 95% confidence interval together with individual data (colored symbols). F values, P values, eta-squared effect size (n^2), and standard error of measurement (SEM) were obtained with a repeated measures ANOVA. #Significantly different from GE_{LAST}, $P \leq 0.001$; $Significantly different from CP, $P \leq 0.030$

demonstrated considerably higher typical errors versus the $7+Y_{LIN}$, $7\text{-}Y_{LIN}$, and GE_{LAST} models.

This is, to our knowledge, the first study that provides novel and detailed information on the within-participant (dis)agreement between four different models of estimating AnWC during cycle ergometry exercise. As hypothesized, the GE_{LAST} model generated significantly higher values of AnWC compared to the $7+Y_{LIN}$ and $7\text{-}Y_{LIN}$ models (~12% higher), as well as to the $CP_{3'AO}$ model (~63% higher). The higher AnWC estimates for GE_{LAST}, when compared to $7+Y_{LIN}$ and $7\text{-}Y_{LIN}$, could be explained by the positive Y-intercept in the $7+Y_{LIN}$ and $7\text{-}Y_{LIN}$ regressions (Fig. 2A) and the increasing GE with higher power output (Fig. 2B), which is contrary to the GE_{LAST} model concept that assumes a constant GE. On average, this resulted in a 0.9 percentage point higher GE during the TT for the $7+Y_{LIN}$ and $7\text{-}Y_{LIN}$ models versus the GE_{LAST} model. It is logical to suggest that the use of a Y-intercept value (i.e., resting baseline value of MR) in the linear PO-MR regression could only be justified if it is reasonably aligned with the submaximal stages of exercise. It appears that this is the case for the current study, because the slopes, delta efficiencies, and Y-intercept values were similar for the $7+Y_{LIN}$ and $7\text{-}Y_{LIN}$ models. This finding is contrary to previous findings for other exercise modalities, such as diagonal-stride treadmill roller-skiing and treadmill running, where the inclusion of a Y-intercept lowered the slope of the regression line significantly and resulted in significantly lower estimates of AnC in highly trained participants (Andersson et al. 2021, 2020). Therefore, the exercise modality should be considered when deciding between the inclusion, or exclusion, of a baseline resting value as a Y-intercept in a linear regression model that is used for determining AnC (or AnWC).
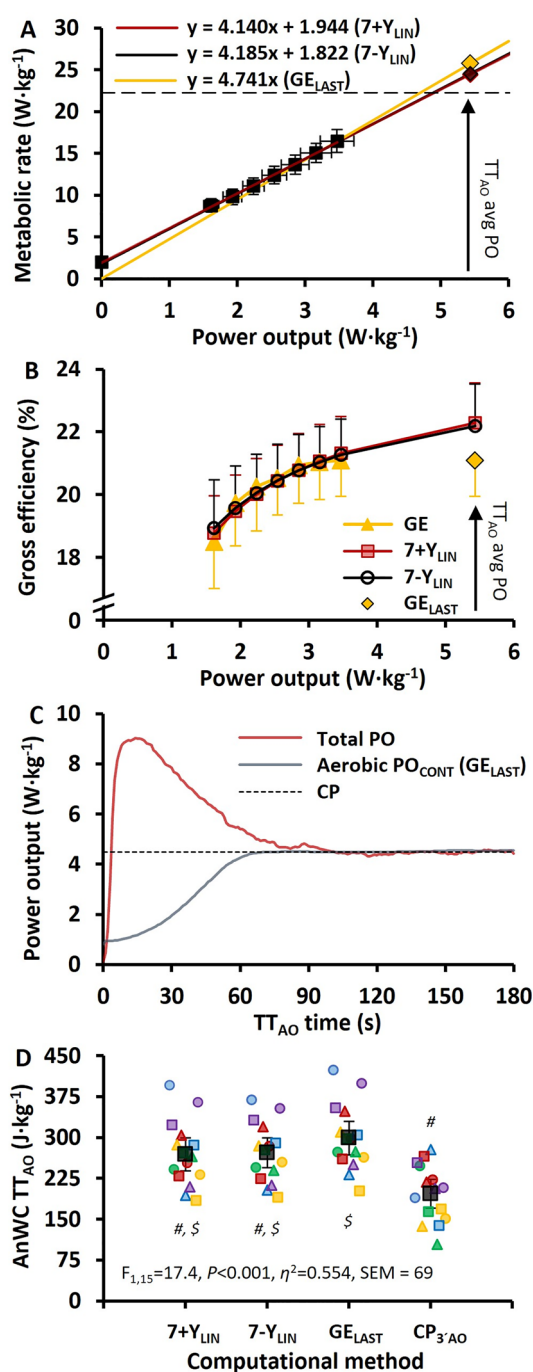
In the current study, the typical errors for the AnWC estimates for the $7+Y_{LIN}$, $7\text{-}Y_{LIN}$, and GE_{LAST} models were relatively low (2–4%). This could be partly explained by the relatively similar within-athlete regressions for the $7+Y_{LIN}$ and $7\text{-}Y_{LIN}$ models (Figs. 3A, B). For instance, the Y-intercept values for the $7+Y_{LIN}$ and $7\text{-}Y_{LIN}$ models were similar (see Figs. 2A and 3A, B). As observed previously (Andersson et al. 2020), the variation in Y-intercept values for the $7\text{-}Y_{LIN}$ model was highly related to the variation in the AnWC estimates between the $7\text{-}Y_{LIN}$ and GE_{LAST}

Table 2 Mean ± standard deviation of slope, delta efficiency, Y-intercept, coefficient of determination (r^2), standard error of estimate (SEE) for the two linear regression models, and gross efficiency, metabolic requirement, and anaerobic capacity during the 3-min all-out cycle time trial for the three different models of estimating the anaerobic capacity

Method of calculation							
	$7+Y_{LIN}$	$7\text{-}Y_{LIN}$	GE_{LAST}	Test statistic	P-value	ES	SEM
Slope (MR per PO [W kg^{-1}])	4.14 ± 0.28	4.19 ± 0.38	–	–	$P = 0.478$	$Hg_{av} = -0.1$	–
Delta efficiency (%)	24.2 ± 1.6	24.1 ± 2.3	–	–	$P = 0.641$	$Hg_{av} = 0.1$	–
Y-intercept (W kg^{-1})	1.94 ± 0.50	1.82 ± 1.00	–	–	$P = 0.449$	$Hg_{av} = 0.1$	–
r^2	0.996 ± 0.003	0.992 ± 0.007	–	–	$P = 0.003$	$Hg_{av} = 0.7$	–
SEE (W kg^{-1})	0.28 ± 0.10	0.25 ± 0.08	–	–	$P = 0.070$	$Hg_{av} = 0.3$	–
GE_{TT_avg} (%)	$22.0 \pm 1.2^*$	$22.0 \pm 1.3^*$	21.1 ± 1.1	$F_{2,28} = 35.9$	$P < 0.001$	$\eta^2 = 0.720$	0.3
MR_{TT_req} (W kg^{-1})	$24.4 \pm 1.8^*$	$24.5 \pm 1.8^*$	25.8 ± 2.0	$F_{2,28} = 50.5$	$P < 0.001$	$\eta^2 = 0.783$	0.4
MR_{TT_req} (% of MR_{ae_peak})	$110 \pm 6^*$	$111 \pm 6^*$	116 ± 7	$F_{2,28} = 48.2$	$P < 0.001$	$\eta^2 = 0.775$	2
AnC (kJ kg^{-1})	$1.17 \pm 0.28^*$	$1.19 \pm 0.24^*$	1.42 ± 0.30	$F_{2,28} = 50.5$	$P < 0.001$	$\eta^2 = 0.783$	0.08

Abbreviations: $7+Y_{LIN}$ and $7\text{-}Y_{LIN}$, the 7×4-min linear models with the baseline metabolic rate as a Y-intercept either included ($7+Y$) or excluded ($7\text{-}Y$), GE_{LAST}, the gross efficiency model based on the last submaximal stage, *MR* metabolic rate, *PO* power output, GE_{TTavg} average GE during the TT, MR_{TTreq} required metabolic rate during the TT, MR_{ae_peak} peak aerobic metabolic rate during the TT, *AnC* anaerobic capacity, Hg_{av} Hedge's g_{av} effect size. *F* values, *P* values, and eta-squared effect size (η^2) were obtained by a one-way ANOVA. *Statistically significantly different from GE_{LAST} ($P < 0.001$)

models (see Fig. 5A). Since the Y-intercept values were all positive and showed a lower between-participants variation compared to previous observations for uphill diagonal-stride roller-skiing (and double poling) (Andersson et al., 2020), the typical errors for the GE_{LAST} versus $7+Y_{LIN}$ and $7\text{-}Y_{LIN}$ models were also considerably lower compared to previous findings for other exercise modalities, such as roller-skiing and running (Andersson et al. 2021, 2020). Even though the typical errors for the GE_{LAST} versus $7+Y_{LIN}$ and $7\text{-}Y_{LIN}$ models were relatively low (typical errors of 2–3%), it is probably still wise not to use different models interchangeably when testing athletes regularly.

Although some previous studies have compared and/or correlated different model estimates of AnWC/AnC derived from cycle-ergometry exercise trials (Dekerle et al. 2006; Green et al. 1994; Hill and Smith 1993; Muniz-Pumares et al. 2016; Noordhof et al. 2011), this is to our knowledge the first study that has compared AnWC estimates based on all the three main methods/concepts for estimating AnWC/AnC in endurance sports (Noordhof et al. 2013). In the current study, the AnWC associated with the $CP_{3'AO}$ model was lower compared to the $7+Y_{LIN}$, $7\text{-}Y_{LIN,}$ and GE_{LAST} models and showed substantial typical errors in AnWC versus the $7+Y_{LIN}$, $7\text{-}Y_{LIN,}$ and GE_{LAST} models (see Figs. 4A–F). Further, the considerably lower AnWC estimate for the $CP_{3'AO}$ model could be explained by the model assumption with CP representing a fixed (i.e., non-dynamic) aerobic contribution to PO and the similar aerobically attributable PO and CP during approximately the second half of TT_{AO} (see Figs. 2C,D). A similar result was found by Dekerle et al. (2006) for a 90-s all-out cycle test when comparing AnWC estimates based on a linear $\dot{v}o_2$-PO regression model with W'

(i.e., "AnWC") based on a conventional CP concept. Some previous studies have shown W', based on any of the CP concepts, to be correlated with other method-specific estimates of AnWC/AnC, which suggests that W' is likely to represent an AnWC (Dekerle et al. 2006; Green et al. 1994; Hill and Smith 1993). However, Muniz-Pumares et al. (2016) compared AnC (expressed as an oxygen deficit) based on the MAOD method with the W' measure (i.e., "AnWC") based on CP derived from the 30-s end power of a 3-min TT_{AO}. The results showed that even if the AnC and W' measures were correlated ($r = 0.654$), there was still a poor agreement between the two measures, suggesting that one, or both, of the methods, was/were unable to accurately measure AnWC/AnC. Thus, both the study of Muniz-Pumares et al. (2016) and the results of the current study question both the validity and reliability of W' derived from a 3-min TT_{AO} as a measure of AnWC, that it was originally considered to represent (Hill 1993; Morton 2006).

Based on the average 30-s end power of TT_{AO}, the CP was found to be 4.5 W kg^{-1}, which for the $7+Y_{LIN}$, $7\text{-}Y_{LIN}$, and GE_{LAST} models represented ~94% of the $\dot{v}o_{2peak}$ reached in TT_{AO}. The current study was not designed to evaluate if the determined CP represented the maximal lactate (or metabolic) steady-state. However, based on previous research findings (Bartram et al. 2017; Iannetta et al. 2020; Karsten et al. 2014; Mattioni Maturana et al. 2016; Sperlich et al. 2011) and that the estimated exercise intensity of CP was 94% of $\dot{v}o_{2peak}$, it is likely that the CP was higher than the maximal lactate (or metabolic) steady state. This indicates that the 3-min TT_{AO} overestimated CP and, as a result, underestimated W', i.e., the "AnWC" parameter, most likely also with questionable reliability
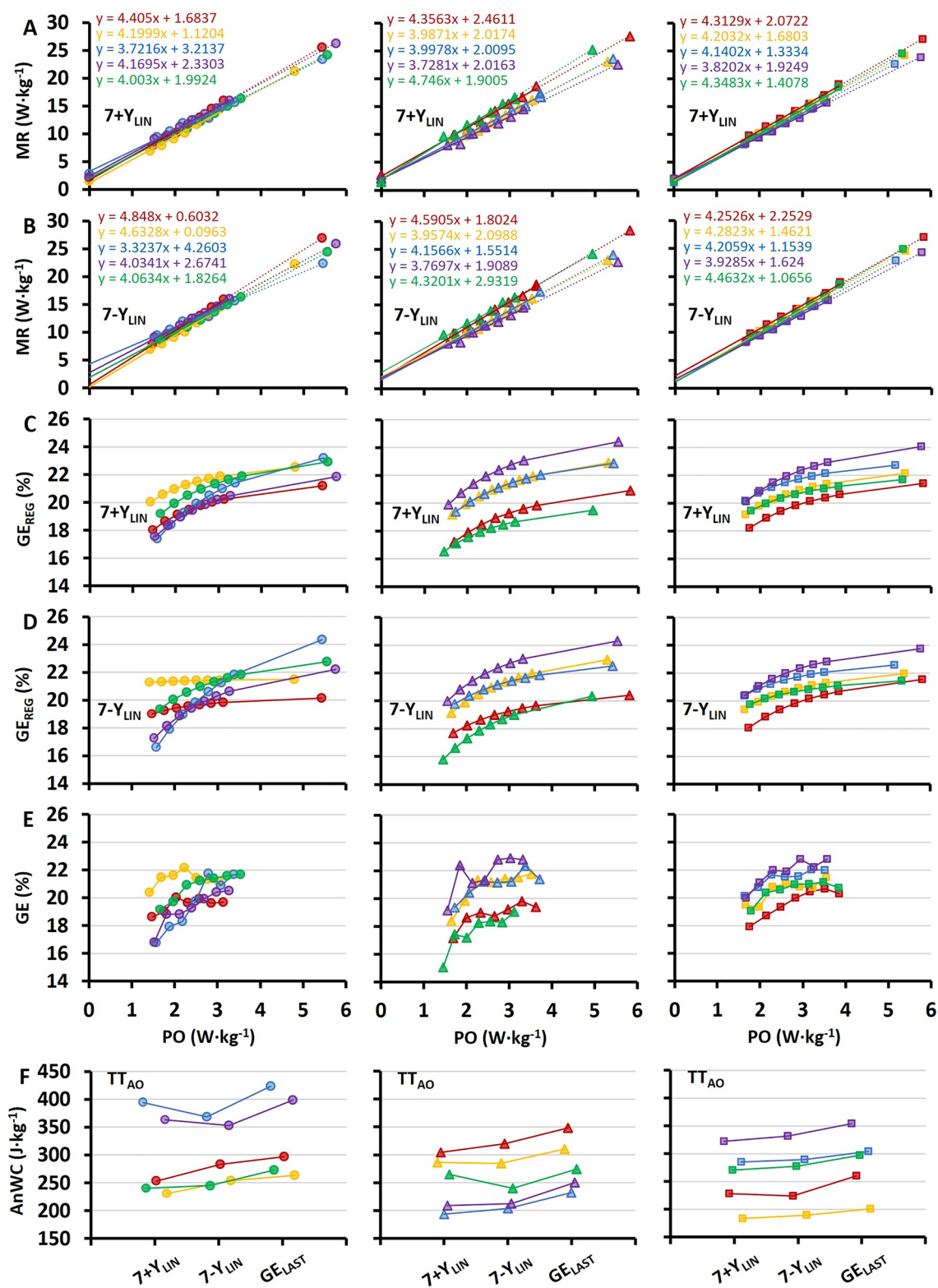
(Bartram et al. 2017; Karsten et al. 2014). This could explain, at least in part, some of the substantial mean differences in AnWC between the $CP_{3'AO}$ model versus the $7 + Y_{LIN}$, $7-Y_{LIN}$, and GE_{LAST} models that were observed in the current study. These results suggest that the *W'* estimate determined from the $CP_{3'AO}$ model is likely to be less valid and reliable as an AnWC compared to the $7 + Y_{LIN}$, $7-Y_{LIN}$, and GE_{LAST} models.

◀**Fig. 3** Individual data ($N=15$, 5 in each of the three horizontal panels) based on 7×4-min of submaximal exercise followed by a 3-min supramaximal all-out TT (TT_{AO}). Regressions for metabolic rate plotted against cycling power output (PO) based on the 7×4-min submaximal stages and the extrapolation up to the average PO during TT_{AO} including a Y-intercept value, i.e., baseline metabolic rate, ($7+Y_{LIN}$) (**A**) and when excluding a Y-intercept value (7-Y_{LIN}) in the respective regressions (**B**). Gross efficiency calculated from the two linear regression equations (GE_{REG}) for the submaximal stages and the TT, with values from $7+Y_{LIN}$ in C and values from 7-Y_{LIN} in **D**. Directly measured values of GE based on the seven submaximal stages (**E**). Individual values of anaerobic work capacity (AnWC) calculated with the three different methods (F), where the $7+Y_{LIN}$ and 7-Y_{LIN} are the two linear models, and the GE_{LAST} model is based on the GE value from the last submaximal stage

It is important to bear in mind that *W'*, based on the $CP_{3'AO}$, and AnWC, as based on any of the three other models (i.e., $7+Y_{LIN}$, 7-Y_{LIN}, and GE_{LAST}), entails quite different methodological concepts. An important factor that is likely to explain the (dis)agreement between *W'* (i.e., "AnWC") based on the $CP_{3'AO}$ model and the other three models is rooted in the aerobic component of *W'*, which arises when aerobically attributable PO is in the intensity domain between CP and the maximal aerobically attributable PO (Vinetti et al. 2017; Vinetti et al. 2017). For example, if the average fractional utilization of $\dot{v}o_{2peak}$ (as a percentage of $\dot{v}o_{2peak}$) during a maximal effort is lower than the corresponding $\dot{v}o_2$ requirement at CP (as a percentage of $\dot{v}o_{2peak}$), *W'* would in theory attain a lower value than AnWC based on the GE_{LAST} model. This was the case in the current study and was partly caused by the overestimated CP. For similar average values of AnWC between the $CP_{3'AO}$ model versus the GE_{LAST} model, the $\dot{v}o_2$ requirement at CP would have needed to be ~80% of $\dot{v}o_{2peak}$ because the average fractional utilization was ~80% of $\dot{v}o_{2peak}$ during TT_{AO}.

As shown in Fig. 2C, the total PO and aerobic contribution to PO (calculated based on GE_{LAST}) were approximately the same after 90 s of TT_{AO}, which indicates that the AnWC/*W'* was depleted after ~90 s with a following plateau of PO during the later stages of the TT, similar to previous studies (Burnley et al. 2006; Vanhatalo et al. 2007). However, a potentially problematic aspect of the 3-min all-out TT concept when using 30-s average end power as a CP measure is related to the fundamental "non-pacing" characteristics, as it is suggested to be a consistent 3-min all-out effort, i.e., with no pacing involved (Burnley et al. 2006). This might be problematic and, thus, susceptible to an increased between-participants variation and decreased reliability, and/or validity, of 30-s end power as a measure of CP. In turn, this would decrease the reliability, and/or validity, of the $CP_{3'AO}$ model for estimating *W'* (i.e., "AnWC") (Bartram et al. 2017; Dotan 2022).

One factor that could have contributed to the overestimated CP during the 3-min TT_{AO} is the different bike settings that were used in the current study compared to the reference study by Burnley et al. (2006). In the current study, cadence and pedal torque were self-selected based on the use of normal shifters, whereas Burnley et al. (2006) used a linear factor set-up on the Lode ergometer (linear factor = power/cadence²), which means that torque increases linearly with cadence and that PO increases quadratically with cadence. The use of a linear factor could, potentially, trigger more of a true all-out maximal effort as PO increases quadratically with cadence and may result in a higher initial PO along with higher exhaustion in the second half of the TT and, thus, a more realistic 30-s end test CP. However, the self-selected cadence approach that was used in the current study resulted in a reasonable cadence for supramaximal exercise and it was only slightly higher than the cadence of 80–90 rev·min⁻¹ that was used in the study by Burnley et al. (2006).

All models of estimating AnWC that are presented in the current study are based on some GE assumptions during exercise at a maximal effort. The GE_{LAST} model assumes a constant GE and the $7+Y_{LIN}$ and 7-Y_{LIN} models assume either an increasing, a constant, or a decreasing GE with the given direction being set by the Y-intercept value in the linear regression (Andersson et al. 2020). Although the $CP_{3'AO}$ is not directly related to GE per se, the model is likely to assume a constant GE since CP is a constant entity, i.e., based on the assumptions that CP is the product of fixed values of aerobic MR and GE, or that aerobic MR and GE change inversely and proportionally so that CP remains constant. One apparent difference between the $7+Y_{LIN}$, 7-Y_{LIN}, and GE_{LAST} models is presented in Fig. 2B, where the $7+Y_{LIN}$ and 7-Y_{LIN} models assume a slightly increasing GE with higher power outputs, whereas the GE_{LAST} model assumes a constant GE during the TT as based on the GE value from the last submaximal stage. Since some previous research findings show that GE declines during high-intensity exercise (de Koning et al. 2013; Noordhof et al. 2015; Sahlin et al. 2005), it is likely that both the $7+Y_{LIN}$ and 7-Y_{LIN} models overestimated GE and consequently underestimated the AnWC/AnC. Therefore, the GE_{LAST} model might have generated a more valid measure of AnWC/AnC. This is also congruous with the results displayed in Fig. 2C and the all-out pacing concept with an early depletion of AnWC (Vanhatalo et al. 2007) and an approximately zero anaerobic power contribution (based on both GE_{LAST} and $CP_{3'AO}$) during the second half of the TT. Although the decline in GE during supramaximal exercise is problematic to measure and includes several assumptions (de Koning et al. 2013), it is very likely that GE declines during high-intensity submaximal and/or supramaximal cycle exercise. This is probably related to a combination of factors, such as exercise hyperpnea (Dempsey et al. 1996),
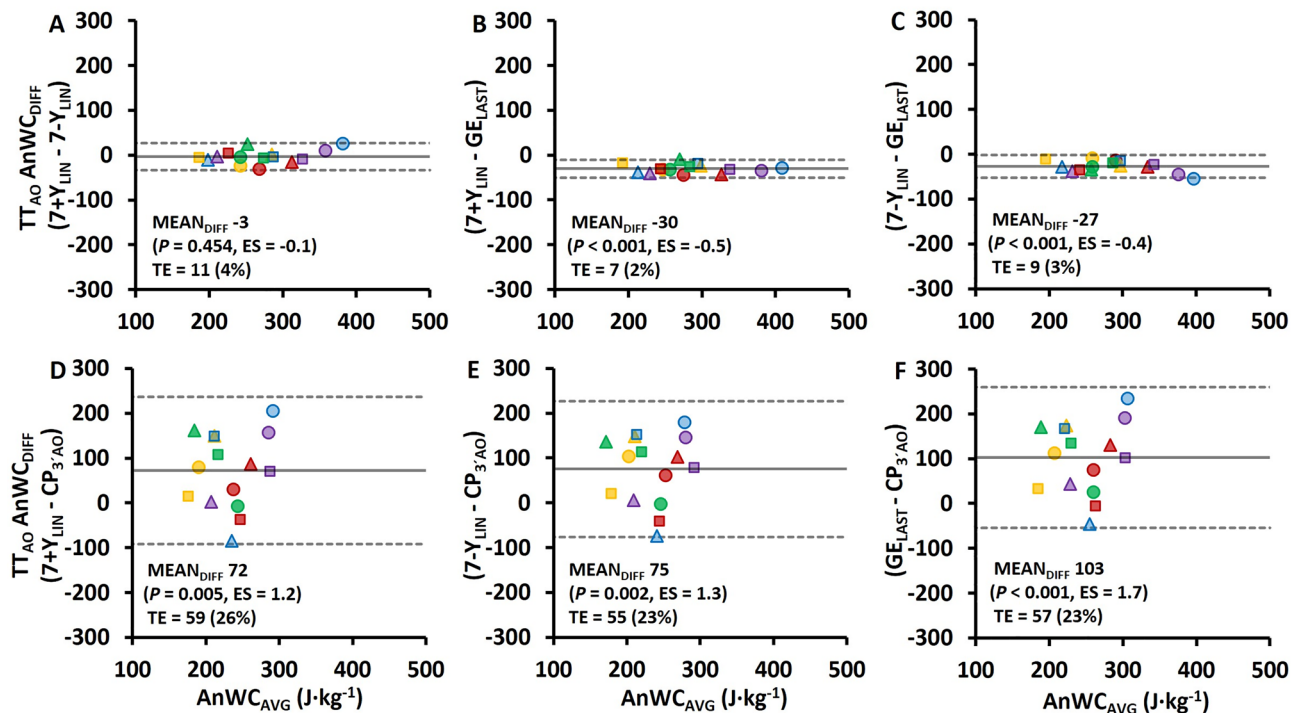
Fig. 4 Bland–Altman plots for the four various models of estimating anaerobic work capacity (AnWC) associated with the 3-min all-out cycle time trial (TT_{AO}) in **A**–**F**. Bland–Altman plots represent the mean difference ($MEAN_{DIFF}$) in the AnWC±95% (1.96 standard deviations) limits of agreement between the methods. Abbreviations: $AnWC_{DIFF}$, the difference in AnWC; TE, absolute typical error (in parenthesis: typical error expressed as a percentage of the grand mean); ES, Hedges's g_{av} effect size (Hg_{av}), $7+Y_{LIN}$ and $7\text{-}Y_{LIN}$, the 7×4-min linear regression methods with the baseline metabolic rate as a Y-intercept either included ($7+Y$) or excluded ($7\text{-}Y$); GE_{LAST}, the gross efficiency model based on the last submaximal stage; $CP_{3'AO}$, the critical power model based on the average 30-s end power during TT_{AO} as critical power. The same symbols used to illustrate individual data in Fig. 3 have also been used in this figure
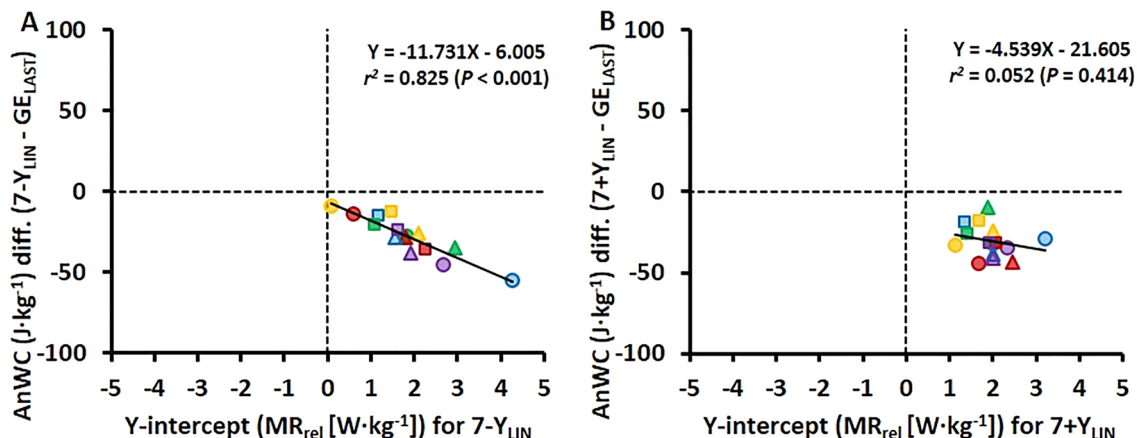
Fig. 5 Scatter plots between the Y-intercept values for the 7×4-min linear regression models with the baseline metabolic rate (MR) as a Y-intercept either excluded ($7\text{-}Y_{LIN}$) or included ($7+Y_{LIN}$) in the model (x-axis) and the anaerobic work capacity difference (AnWC diff.) versus the gross efficiency method based on the last submaximal stage (GE_{LAST}) (y-axis)

altered muscle recruitment patterns, and/or fatigue (Sahlin et al. 2005). All these factors are likely to explain most of the $\dot{v}o_2$ slow component during high- and supramaximal-intensity exercise (Sahlin et al. 2005). When considering all these factors, the most valid methodology for determining AnC/AnWC during cycle ergometry exercise for a group with relatively homogenous cardiovascular fitness (i.e., $\dot{v}o_{2max}$) is likely to be GE measured at a relatively

high submaximal exercise intensity ($\sim 75\%$ of $\dot{v}o_{2max}$). An alternative solution could be to use the $7 + Y_{LIN}$ or $7\text{-}Y_{LIN}$ model and determine a fixed GE value calculated from the PO-MR regression based on a MR that corresponds to $\sim 75\%$ of the maximal aerobic MR (i.e., $\dot{v}o_{2max}$ expressed as a MR). This GE value can then be used similarly to the GE_{LAST} model for determining AnC/AnWC.

This study provides practical information for exercise physiologists by demonstrating the (dis)agreement between different models of estimating AnWC/AnC. Exercise physiologists should be aware that the use of these models is not interchangeable. The optimal model is likely to be, at least to some extent, case-by-case dependent, however, this study suggests that particular caution should be used when interpreting W' from the $CP_{3'AO}$ method as a measure of AnWC. As the test-to-test reliability of AnWC/ AnC is of high practical importance to athletes/coaches, future research is needed to determine the test-to-test reliability of different models used for estimating AnWC/AnC.

In conclusion, the $7 + Y_{LIN}$ and $7\text{-}Y_{LIN}$ models generated 10% lower AnWC values than the GE_{LAST} model. This result was caused by the ~ 0.9 percentage points higher supramaximal GE for the $7 + Y_{LIN}$ and $7\text{-}Y_{LIN}$ (calculated based on the linear PO-MR regression). When expressed as an AnC, the same comparison generated 17% lower values for the $7 + Y_{LIN}$ and $7\text{-}Y_{LIN}$ models versus the GE_{LAST} model. Due to the similar slopes and Y-intercepts for the $7 + Y_{LIN}$ and $7\text{-}Y_{LIN}$ models, supramaximal GE and AnWC/ AnC were the same for both models. The within-participants variation in AnWC estimates between the $7 + Y_{LIN}$, $7\text{-}Y_{LIN}$, and GE_{LAST} models was low, as indicated by the low typical errors that ranged between 7 and 11 J•kg^{-1} (or 2–4% of the respective pair-wise grand mean). The $CP_{3'AO}$ model generated the least valid and reliable estimate of AnWC, as revealed by the 30% lower AnWC values (on average) compared to the other models, and the substantial between-models typical errors (55–59 J•kg^{-1}, or 23–26% of the respective pair-wise grand mean).

References

Andersson EP, Noordhof DA, Lögdal N (2020) The anaerobic capacity of cross-country skiers: the effect of computational method and skiing sub-technique. Front Sports Act Living. https://doi.org/10.3389/fspor.2020.00037

Andersson EP, Björklund G, McGawley K (2021) Anaerobic capacity in running: the effect of computational method. Front Physiol. https://doi.org/10.3389/fphys.2021.708172

Andersson EP, McGawley K (2018) A comparison between different methods of estimating anaerobic energy production. Front Physiol 9:82. https://doi.org/10.3389/fphys.2018.00082

Bartram JC, Thewlis D, Martin DT, Norton KI (2017) Predicting critical power in elite cyclists: questioning the validity of the 3-minute all-out test. Int J Sports Physiol Perform 12(6):783–787. https://doi.org/10.1123/ijspp.2016-0376

Bland JM, Altman DG (1999) Measuring agreement in method comparison studies. Stat Methods Med Res 8(2):135–160. https://doi.org/10.1177/096228029900800204

Borg GA (1982) Psychophysical bases of perceived exertion. Med Sci Sports Exerc 14(5):377–381

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

de Koning JJ, Noordhof DA, Uitslag TP, Galiart RE, Dodge C, Foster C (2013) An approach to estimating gross efficiency during high-intensity exercise. Int J Sports Physiol Perform 8(6):682–684. https://doi.org/10.1123/ijspp.8.6.682

Dekerle J, Brickley G, Hammond AJ, Pringle JS, Carter H (2006) Validity of the two-parameter model in estimating the anaerobic work capacity. Eur J Appl Physiol 96(3):257–264. https://doi.org/10.1007/s00421-005-0074-8

Dempsey JA, Harms CA, Ainsworth DM (1996) Respiratory muscle perfusion and energetics during exercise. Med Sci Sports Exerc 28(9):1123–1128. https://doi.org/10.1097/00005768-199609000-00007

Dotan R (2022) A critical review of critical power. Eur J Appl Physiol. https://doi.org/10.1007/s00421-022-04922-6

Ettema G, Lorås HW (2009) Efficiency in cycling: a review. Eur J Appl Physiol 106(1):1–14. https://doi.org/10.1007/s00421-009-1008-7

Green S, Dawson BT, Goodman C, Carey MF (1994) Y-intercept of the maximal work-duration relationship and anaerobic capacity in cyclists. Eur J Appl Physiol Occup Physiol 69(6):550–556. https://doi.org/10.1007/bf00239874

Hill DW (1993) The critical power concept. A Review Sports Med 16(4):237–254. https://doi.org/10.2165/00007256-199316040-00003

Hill DW, Smith JC (1993) A comparison of methods of estimating anaerobic work capacity. Ergonomics 36(12):1495–1500. https://doi.org/10.1080/00140139308968017

Iannetta D, Inglis EC, Mattu AT, Fontana FY, Pogliaghi S, Keir DA, Murias JM (2020) A critical evaluation of current methods for exercise prescription in women and men. Med Sci Sports Exerc 52(2):466–473. https://doi.org/10.1249/mss.0000000000002147

Karsten B, Jobson SA, Hopker J, Passfield L, Beedie C (2014) The 3-min test does not provide a valid measure of critical power using the SRM isokinetic mode. Int J Sports Med 35(4):304–309. https://doi.org/10.1055/s-0033-1349093

Lakens D (2013) Calculating and reporting effect sizes to facilitate cumulative science: a practical primer for t-tests and ANOVAs. Front Psychol 4:863. https://doi.org/10.3389/fpsyg.2013.00863

Mattioni Maturana F, Keir DA, McLay KM, Murias JM (2016) Can measures of critical power precisely estimate the maximal metabolic steady-state? Appl Physiol Nutr Metab 41(11):1197–1203. https://doi.org/10.1139/apnm-2016-0248

McKay AKA, Stellingwerff T, Smith ES, Martin DT, Mujika I, Goosey-Tolfrey VL, Sheppard J, Burke LM (2022) Defining training and performance caliber: a participant classification framework. Int J Sports Physiol Perform 17(2):317–331. https://doi.org/10.1123/ijspp.2021-0451

Medbø JI, Mohn AC, Tabata I, Bahr R, Vaage O, Sejersted OM (1988) Anaerobic capacity determined by maximal accumulated O2 deficit. J Appl Physiol 64(1):50–60. https://doi.org/10.1152/jappl.1988.64.1.50

Morton RH (2006) The critical power and related whole-body bioenergetic models. Eur J Appl Physiol 96(4):339–354. https://doi.org/10.1007/s00421-005-0088-2

Muniz-Pumares D, Pedlar C, Godfrey R, Glaister M (2016) A comparison of methods to estimate anaerobic capacity: accumulated oxygen deficit and W' during constant and all-out work-rate profiles. J Sports Sci. https://doi.org/10.1080/02640414.2016.1267386

Nieman DC, Austin MD, Dew D, Utter AC (2013) Validity of COSMED's quark CPET mixing chamber system in evaluating energy metabolism during aerobic exercise in healthy male adults. Res Sports Med 21(2):136–145. https://doi.org/10.1080/15438627.2012.757227

Noordhof DA, de Koning JJ, Foster C (2010) The maximal accumulated oxygen deficit method: a valid and reliable measure of anaerobic capacity? Sports Med 40(4):285–302. https://doi.org/10.2165/11530390-000000000-00000

Noordhof DA, Vink AM, de Koning JJ, Foster C (2011) Anaerobic capacity: effect of computational method. Int J Sports Med 32(6):422–428. https://doi.org/10.1055/s-0031-1271676

Noordhof DA, Skiba PF, de Koning JJ (2013) Determining anaerobic capacity in sporting activities. Int J Sports Physiol Perform 8(5):475–482. https://doi.org/10.1123/ijspp.8.5.475

Noordhof DA, Mulder RC, Malterer KR, Foster C, de Koning JJ (2015) The decline in gross efficiency in relation to cycling time-trial length. Int J Sports Physiol Perform 10(1):64–70. https://doi.org/10.1123/ijspp.2014-0034

Sahlin K, Sorensen JB, Gladden LB, Rossiter HB, Pedersen PK (2005) Prior heavy exercise eliminates VO2 slow component and reduces efficiency during submaximal exercise in humans. J Physiol 564(Pt 3):765–773. https://doi.org/10.1113/jphysiol.2005.083840

Serresse O, Lortie G, Bouchard C, Boulay MR (1988) Estimation of the contribution of the various energy systems during maximal work of short duration. Int J Sports Med 9(6):456–460. https://doi.org/10.1055/s-2007-1025051

Shaw AJ, Ingham SA, Folland JP (2014) The valid measurement of running economy in runners. Med Sci Sports Exerc 46(10):1968–1973. https://doi.org/10.1249/mss.0000000000000311

Sperlich B, Haegele M, Thissen A, Mester J, Holmberg HC (2011) 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(6):433–437. https://doi.org/10.1055/s-0031-1271770

Vanhatalo A, Doust JH, Burnley M (2007) Determination of critical power using a 3-min all-out cycling test. Med Sci Sports Exerc 39(3):548–555. https://doi.org/10.1249/mss.0b013e31802dd3e6

Vanhatalo A, Jones AM, Burnley M (2011) Application of critical power in sport. Int J Sports Physiol Perform 6(1):128–136. https://doi.org/10.1123/ijspp.6.1.128

Vinetti G, Fagoni N, Taboni A, Camelio S, di Prampero PE, Ferretti G (2017) Effects of recovery interval duration on the parameters of the critical power model for incremental exercise. Eur J Appl Physiol 117(9):1859–1867. https://doi.org/10.1007/s00421-017-3662-5

Vinetti G, Taboni A, Bruseghini P, Camelio S, D'Elia M, Fagoni N, Moia C, Ferretti G (2019) Experimental validation of the 3-parameter critical power model in cycling. Eur J Appl Physiol 119(4):941–949. https://doi.org/10.1007/s00421-019-04083-z

Weir JB (1949) New methods for calculating metabolic rate with special reference to protein metabolism. J Physiol 109(1–2):1–9. https://doi.org/10.1113/jphysiol.1949.sp004363

Winkert K, Kirsten J, Dreyhaupt J, Steinacker JM, Treff G (2020) The COSMED K5 in breath-by-breath and mixing chamber mode at low to high intensities. Med Sci Sports Exerc 52(5):1153–1162. https://doi.org/10.1249/mss.0000000000002241

Winter EM, Fowler N (2009) Exercise defined and quantified according to the Systeme International d'Unites. J Sports Sci 27(5):447–460. https://doi.org/10.1080/02640410802658461