



# A longitudinal study on the interchangeable use of whole-body and local exercise thresholds in cycling

Kevin Caen<sup>1,2</sup> · Jan G. Bourgois<sup>1,2</sup> · Eva Stassijns<sup>1</sup> · Jan Boone<sup>1,2</sup>

Received: 9 August 2021 / Accepted: 23 March 2022 / Published online: 18 April 2022  
© The Author(s), under exclusive licence to Springer-Verlag GmbH Germany, part of Springer Nature 2022

## Abstract

**Purpose** This study longitudinally examined the interchangeable use of critical power (CP), the maximal lactate steady state (MLSS) and the respiratory compensation point (RCP) (i.e., whole-body thresholds), and breakpoints in muscle deoxygenation ( $m[\text{HHb}]_{\text{BP}}$ ) and muscle activity ( $i\text{EMG}_{\text{BP}}$ ) (i.e., local thresholds).

**Methods** Twenty-one participants were tested on two timepoints (T1 and T2) with a 4-week period (study 1: 10 women, age =  $27 \pm 3$  years,  $\dot{V}\text{O}_{2\text{peak}} = 43.2 \pm 7.3$  mL min<sup>-1</sup> kg<sup>-1</sup>) or a 12-week period (study 2: 11 men, age =  $25 \pm 4$  years,  $\dot{V}\text{O}_{2\text{peak}} = 47.7 \pm 5.9$  mL min<sup>-1</sup> kg<sup>-1</sup>) in between. The test battery included one ramp incremental test (to determine RCP,  $m[\text{HHb}]_{\text{BP}}$  and  $i\text{EMG}_{\text{BP}}$ ) and a series of (sub)maximal constant load tests (to determine CP and MLSS). All thresholds were expressed as oxygen uptake ( $\dot{V}\text{O}_2$ ) and equivalent power output (PO) for comparison.

**Results** None of the thresholds were significantly different in study 1 ( $\dot{V}\text{O}_2$ :  $P = 0.143$ , PO:  $P = 0.281$ ), but differences between whole-body and local thresholds were observed in study 2 ( $\dot{V}\text{O}_2$ :  $P < 0.001$ , PO:  $P = 0.024$ ). Whole-body thresholds showed better 4-week test–retest reliability (TEM = 88–125 mL min<sup>-1</sup> or 6–10 W, ICC = 0.94–0.98) compared to local thresholds (TEM = 189–195 mL min<sup>-1</sup> or 15–18 W, ICC = 0.58–0.89). All five thresholds were strongly associated at T1 and T2 ( $r = 0.75$ – $0.99$ ), but their changes from T1 to T2 were mostly uncorrelated ( $r = -0.41$ – $0.83$ ).

**Conclusion** Whole-body thresholds (CP/MLSS/RCP) showed a close and consistent coherence taking into account a 3–6%-bandwidth of typical variation. In contrast, local thresholds ( $m[\text{HHb}]_{\text{BP}}/i\text{EMG}_{\text{BP}}$ ) were characterized by higher variability and did not consistently coincide with the whole-body thresholds. In addition, we found that most thresholds evolved independently of each other over time. Together, these results do not justify the interchangeable use of whole-body and local exercise thresholds in practice.

**Keywords** Critical power · Maximal lactate steady state · Respiratory compensation point · Deoxygenation threshold · EMG threshold

## Abbreviations

|                   |                                    |
|-------------------|------------------------------------|
| CI <sub>95%</sub> | 95% Confidence interval            |
| CP                | Critical power                     |
| CV%               | Coefficient of variation           |
| La <sup>-</sup>   | Lactate                            |
| GET               | Gas exchange threshold             |
| ICC               | Intraclass correlation coefficient |

|                             |   |
|-----------------------------|---|
| $i\text{EMG}_{\text{BP}}$   | Breakpoint in the integrated electromyography signal                |
| MD                          | Minimum difference  |
| $m[\text{HHb}]_{\text{BP}}$ | Breakpoint in muscle deoxygenated hemo- and myoglobin concentration |
| MLSS                        | Maximal lactate steady state  |
| MMSS                        | Maximal metabolic steady state                                      |
| MRT                         | Mean response time  |
| NIRS                        | Near-infrared spectroscopy  |
| PO                          | Power output  |
| RCP                         | Respiratory compensation point                                      |
| TEM                         | Typical error of measurement  |
| $\dot{V}\text{O}_2$         | Pulmonary oxygen uptake   |

Communicated by Philip D. Chilibeck.

✉ Jan Boone  
Jan.Boone@UGent.be

<sup>1</sup> Department of Movement and Sports Sciences, Ghent University, Watersportlaan 2, 9000 Ghent, Belgium

<sup>2</sup> Center of Sports Medicine, Ghent University Hospital, Ghent, Belgium

## Introduction

Based on the physiological response to exercise, three distinct domains of exercise intensity have been defined (i.e., moderate, heavy and severe) (Poole et al. 1988; Burnley and Jones 2007). The threshold between heavy and severe exercise intensity is referred to as the maximal metabolic steady-state (MMSS) intensity (Jones et al. 2019). As this threshold distinguishes exercise intensities for which the intramuscular metabolic milieu can(not) be stabilized, it is a commonly used index for testing, training and monitoring athletes. Traditionally, two threshold concepts have been considered as reference methods to determine the MMSS intensity: critical power, which defines the hyperbolic relationship between power output (PO) and time to exhaustion (Poole et al. 2016), and the maximal lactate ( $\text{La}^-$ ) steady state (MLSS) (Billat et al. 2003; Faude et al. 2009), which defines the highest possible equilibrium in blood [ $\text{La}^-$ ] during exercise.

Due to the time-consuming methodology of CP and MLSS, sports practitioners and clinicians often estimate the MMSS intensity from a single ramp incremental exercise test. During such test, gas exchange measures ( $\dot{V}\text{O}_2$ ,  $\dot{V}\text{CO}_2$ , and  $\dot{V}\text{E}$ ) allow for the detection of the respiratory compensation point (RCP). This point denotes the onset of hyperventilation engendered by metabolic acidosis when crossing the heavy-to-severe exercise boundary (Whipp et al. 1989). Just like CP and MLSS, RCP is considered as a whole-body threshold because it reflects the generic physiological response of the body to exercise. In contrast to these whole-body thresholds, also more local exercise thresholds (i.e., measured directly at the level of the working muscles) have been identified (Boone et al. 2016). A first local threshold can be detected using near-infrared spectroscopy (NIRS) and is manifested as a plateau-like response in the signal of deoxygenated hemo- and myoglobin ( $\text{m}[\text{HHb}]_{\text{BP}}$ ), a parameter that is regarded as an index of local muscle  $\text{O}_2$  extraction (Spencer et al. 2012). This breakpoint has been linked to an increased  $\text{O}_2$  provision, likely related to metabolic vasodilation and blood flow redistribution, and may indicate the attainment of an upper limit in muscle  $\text{O}_2$  extraction and a progressively larger contribution from non-oxidative energy turnover towards the end of a ramp test (Bellotti et al. 2013; Fontana et al. 2015). A second local threshold can be detected using surface electromyography (EMG) and exhibits a nonlinear increase in the integrated EMG signal ( $\text{iEMG}_{\text{BP}}$ ) (Nagata et al. 1981). This increase in muscle activity is related to the additional recruitment of mainly type-II muscle fibers with increasing intensity during ramp exercise (Takaishi et al. 1992).

Despite the fact that whole-body and local exercise thresholds occur within a narrow range of exercise intensities

(~75–90%  $\dot{V}\text{O}_{2\text{max}}$ ), and that they may be mechanically linked (Boone et al. 2016), the available research gives opposing viewpoints on their mutual agreement, as well as an on the presumption that each of these thresholds represents the actual MMSS intensity (Craig et al. 2015; Keir et al. 2015a, 2018b, 2019; Broxterman et al. 2018; Marwood et al. 2019). Part of the contradiction originates from a lack of methodological consistency among the existing studies (i.e., differences in test protocols and threshold determination). In addition, it should be noted that the vast majority of these studies are cross-sectional (see Caen (2021) for an overview table of previous studies). To date, only two studies have compared the interrelationship of whole-body and local exercise thresholds in a longitudinal way. Caen et al. (2018) evaluated the interrelationship of CP, RCP and  $\text{m}[\text{HHb}]_{\text{BP}}$  in untrained men before and after a 6-week training intervention, whereas Inglis et al. (2020) compared MLSS, RCP and  $\text{m}[\text{HHb}]_{\text{BP}}$  in a group of trained cyclists across a 7-month cycling season. Unfortunately, both studies came to different conclusions, which again contributed to the already existing contrasts in the literature.

New studies to investigate whole-body and local exercise thresholds longitudinally are warranted for several reasons. First, it remains difficult to gain full insight into the interrelationship of CP, MLSS, RCP,  $\text{m}[\text{HHb}]_{\text{BP}}$  and  $\text{iEMG}_{\text{BP}}$  as these thresholds have never been studied all together. Indeed, previous studies have always focused on few thresholds at once. For example, the studies of Caen et al. (2018) and Inglis et al. (2020) included only one criterion method, either CP or MLSS, against which the others were compared. A longitudinal comparison including both indices may further elucidate their coincidence. Second, studies providing comprehensive data on the reproducibility of exercise thresholds are scarce, although such information would be of great importance for scientists and coaches to better interpret the significance of potential differences between them (i.e., natural variation vs. actual changes). Currently, there is limited information on the typical variation expected when thresholds are determined over a short period of time. At last, from a practical point of view, threshold equivalence would mean that all thresholds could be used interchangeably in the field. However, as previously stated by Keir et al. (2018b), this is only feasible when these thresholds can be accurately expressed in terms of PO (i.e., the main parameter used for cycling training). So far, it has been recommended to compare thresholds only by their associated metabolic rate ( $\dot{V}\text{O}_2$ ) and not in terms of PO. The rationale for this is that PO values of thresholds obtained from ramp incremental exercise (e.g., RCP and local thresholds) are protocol-dependent and elicit a higher metabolic load than intended. This latter issue is related to the discrepancy in the  $\dot{V}\text{O}_2/\text{PO}$  relationships between ramp and constant load exercise (Keir

et al. 2018a). Only recently, strategies to eliminate this “gap” were developed and successfully applied to translate RCP into its equivalent PO for exercise prescription (Caen et al. 2020; Iannetta et al. 2020). However, at this point, no study has applied these corrections to compare whole-body and local exercise thresholds. Currently, their interchangeable use in practice remains therefore unknown.

The purpose of this study was to examine whether whole-body and local exercise thresholds can be used interchangeably to estimate the MMSS intensity. To answer this question, we performed two different study designs including three whole-body thresholds (i.e., CP, MLSS and RCP) and two local thresholds (i.e.,  $m[\text{HHb}]_{\text{BP}}$  and  $i\text{EMG}_{\text{BP}}$ ) that are commonly used. In study 1, we examined the threshold interrelationship and evaluated their reproducibility by means of a 4-week test–retest reliability design. In study 2, we investigated the threshold interrelationship by means of a 12-week intervention study. In both studies, all thresholds were compared when expressed as  $\dot{V}\text{O}_2$  and PO.

## Methods

### Participants

Twenty-one young and healthy persons took part in this study on a voluntary basis. The group was split into a female participant group ( $n = 10$ ,  $27 \pm 3$  years,  $1.69 \pm 0.04$  m,  $60.7 \pm 7.5$  kg), which was assigned to study 1, and a male participant group ( $n = 11$ ,  $25 \pm 4$  years,  $1.81 \pm 0.07$  m,  $78.1 \pm 12.5$  kg), which was assigned to study 2. All participants reported to have a physically active lifestyle with sports practice on a regular basis ( $5 \pm 2$  h/week). Subjects underwent a medical screening prior to the start of the study to ensure that they were in good health. Informed consent was received from all participants. This study was conform to the Declaration of Helsinki (1964) and approved by the Ethical Committee of the Ghent University Hospital (Ghent, Belgium).

### Experimental procedure and protocols

All participants performed the same test battery at two distinct timepoints (i.e., T1 and T2), interspersed with either a 4-week period (i.e., study 1, female group) or a 12-week period (i.e., study 2, male group). The test battery included one ramp incremental exercise test to determine RCP,  $m[\text{HHb}]_{\text{BP}}$  and  $i\text{EMG}_{\text{BP}}$ , three to five constant load tests to exhaustion to determine CP, and a minimum of two 30-min constant load tests to determine MLSS.

The female participant group ( $n = 10$ ) was assigned to study 1, which involved a 4-week test–retest reliability study. Between T1 and T2, there was a 4-week “wash-out” period

to counteract possible training effects due to the extensive number of test sessions (i.e., ~6–8 test days). Participants of study 1 were instructed to maintain their habitual level of physical activity, sports practice and diet.

The male participant group ( $n = 11$ ) was assigned to study 2, which involved a 12-week intervention study. The intervention consisted of a nationwide lockdown that was implemented as part of the protective measures against the COVID-19 pandemic. During this lockdown period, sports practice was restricted to workouts at home and a limited amount of outdoor activities (e.g., walking, running, cycling). Participants of study 2 did not receive any instructions with regard to their sports practice or physical activity level. Seven participants reported an increase in weekly physical exercise, three participants reported a decrease, and one participant indicated that he maintained the same level of physical exercise during the lockdown period. On average, weekly training volume (e.g., cycling, running, workouts) was  $5.0 \pm 3.5$  h/week at T1 and  $6.6 \pm 2.5$  h/week at T2.

All exercise tests were performed in the Sport Science Laboratory—Jacques Rogge (Ghent, Belgium) on an electromagnetically braked cycle ergometer (Excalibur Sport, Lode, Groningen, The Netherlands) of which the saddle and handlebar were individually set for each test. Room air temperature was 18 °C and relative humidity was ~50%. During the first test, participants were instructed to cycle at a fixed pedal cadence between 70 and 90 rpm. This self-selected cadence was imposed during all subsequent tests. Lab visits were always planned at a fixed time of the day, with a total of two or three tests per week, but never on two consecutive days. Participants were asked to avoid strenuous exercise in the last 24 h before testing, and to standardize food and drink intake on test days.

During all tests, gas exchange was registered breath-by-breath using a metabolic system (Metalyzer 3B, Cortex, Leipzig, Germany) that was calibrated before each measurement according to the manufacturer’s recommendations. Heart rate (HR) was continuously monitored using a chest strap (H7, Polar, Kempele, Finland).

### Ramp incremental test

During the first visit, participants completed a  $30 \text{ W} \cdot \text{min}^{-1}$  ramp incremental test to exhaustion. The test protocol was modified from Iannetta et al. (2019), starting with a 6-min constant load bout at 20 W, followed by a 6-min constant load bout at 80 W (for females) or 100 W (for males). Then, after 2 min of seated rest, subjects cycled an additional 4 min at 50 W (i.e., baseline cycling) before the continuous increase in PO started. The test was terminated at voluntary exhaustion, which was defined as the inability to maintain

the preferred cadence for more than 5 s despite strong verbal encouragement.

Muscle deoxygenation (by means of NIRS) and muscle activity (by means of sEMG) were measured at the M. Vastus Lateralis of the right and left leg, respectively. Both measuring sites were shaved and cleaned with alcohol before applying the equipment to the skin. Muscle deoxygenation was recorded using a NIRS tissue oximeter (OxiplexTS, ISS, Champaign, IL, USA) at a sampling rate of 25 Hz. The probe was attached longitudinally to the distal portion of the muscle belly and wrapped with a bandage to prevent it from displacing during cycling. Muscle activity was registered using a wireless sEMG system (Telemetry Desktop DTS, Noraxon, Scottsdale, AZ, USA) with disposable, self-adhesive Ag–AgCl electrodes (2 cm spacing). Electrodes were placed in accordance with the SENIAM guidelines (Hermens et al. 2000), and connected with a DTS EMG sensor which transmitted the myoelectric signals to the DTS receiver at a sampling rate of 1500 Hz. Prior to the start of each test, NIRS and sEMG signals were checked for noise and abnormalities.

### CP trials

Subjects completed three constant load tests to exhaustion. Each test started with 4 min of baseline cycling at 50 W, immediately followed by an abrupt increase to the appropriate PO at which the subjects had to cycle as long as possible. PO were chosen on an individual basis between 65 and 100% of the peak PO ( $PO_{peak}$ ) attained during the ramp incremental test, with the aim of eliciting times to exhaustion between 2 and 20 min. To acquire a proportional distribution of times, the objective was to obtain at least one trial of short duration (2–4 min), one trial of medium duration (4–10 min) and one trial of longer duration (12–20 min). If the determination criteria for CP and the work capacity above CP (i.e., termed  $W$ ) (see “Data analysis” section) were not met after three tests, one or more additional trials had to be performed.

### MLSS trials

After 4 min of baseline cycling at 50 W, participants completed a 30-min constant load test at their individually calculated CP. At the end of baseline cycling and from then every 5 min, capillary blood samples (20  $\mu$ L) were drawn from the fingertip to measure  $[La^-]$  (Biosen C-Line, EKF Diagnostics GmbH, Berlin, Germany). Depending on the  $La^-$  response, participants had to perform a total of two or more trials to identify the MLSS intensity (see “Data analysis” section).

## Data analysis

### Ramp incremental test

Raw  $\dot{V}O_2$  data were cleaned by removing all data points that lay outside the 95% prediction bands, linearly interpolated per second and then converted into 10-s values for further analysis (Origin 2020, OriginLab, Northampton, MA, USA).  $\dot{V}O_{2peak}$  and peak respiratory exchange ratio ( $RER_{peak}$ ) were defined as the highest 30-s values at the end of the ramp incremental test.  $HR_{peak}$  was defined as the highest value throughout.

The gas exchange threshold (GET) was identified as the first disproportionate increase in  $\dot{V}CO_2$  relative to  $\dot{V}O_2$  (V-slope method), and was verified by detecting the first increase in the  $\dot{V}E$ ,  $\dot{V}E/\dot{V}O_2$  and  $P_{ET}O_2$  response (Beaver et al. 1986; Binder et al. 2008). RCP was determined by identifying the simultaneous occurrence of the first increase in  $\dot{V}E/\dot{V}CO_2$ , the second increase in  $\dot{V}E$  and the deflection point in  $P_{ET}CO_2$  (Beaver et al. 1986; Binder et al. 2008). GET and RCP were visually determined by three independent experts and consensus was reached for all subjects.

Raw sEMG signals were processed to create a linear envelope by applying rectification, a bandpass filter (10–500 Hz) and a root mean square (RMS) smoothing with a time window of 100 ms (MyoResearch 3.2, Noraxon, Scottsdale, AZ, USA). Subsequently, the total accumulated activity for each second was calculated (iEMG) and signals were converted into 10-bins. Raw NIRS signals included [HHb],  $[O_2Hb]$ , total [Hb] and tissue oxygenation index. Of these indices, only [HHb] data were averaged per 10 s and retained for further analysis.

$m[HHb]_{BP}$  and  $iEMG_{BP}$  were determined using a double-linear regression model by plotting the measured signals against PO, yielding two linear functions (Origin 2020, OriginLab, Northampton, MA, USA):

$$\text{For } x < \text{BP: } y = m_1x + b_1,$$

$$\text{For } x > \text{BP: } y = m_2x + b_2,$$

where  $m$  represents the regression slope,  $b$  corresponds to the y-intercept and BP is the break point calculated by the least-squares model fit. More specifically, the intersection of the two linear segments corresponded to the  $m[HHb]_{BP}$  as the point where there was a flattening of the [HHb] response, and to the  $iEMG_{BP}$  as the point where there was an amplification in the iEMG signal. The fitting window of the regression model was visually adjusted from the onset of the systematic rise in the  $m[HHb]$  or iEMG signal until the end of the test.

For each participant, the individual  $\dot{V}O_2/PO$  relationship was linearly modeled in two portions:

- Portion 1: including all data points between the start of linear rise in  $\dot{V}O_2$  and the uncorrected PO at GET
- Portion 2: including all data points between the uncorrected PO at GET and the end of the linear rise in  $\dot{V}O_2$ , leaving out data points that were related to the plateau in  $\dot{V}O_2$  towards the end of the test

The slopes of the regression lines that characterize portion 1 and 2 reflect the absolute gain in  $\dot{V}O_2$  (expressed in  $\text{mL min}^{-1} \text{W}^{-1}$ ) and are defined as  $s_{1\text{-ramp}}$  and  $s_{2\text{-ramp}}$ , respectively. The equations of portion 1 and 2 were used to calculate the exact  $\dot{V}O_2$  at which GET, RCP,  $m[\text{HHb}]_{\text{BP}}$  and  $i\text{EMG}_{\text{BP}}$  occurred.

The ramp-identified PO at GET was corrected by means of the mean response time (MRT). The MRT was calculated based on the steady-state  $\dot{V}O_2$  at 80 W (for females) or 100 W (for males) (see Iannetta et al. (2019) for a detailed description), and expressed in W. Subsequently, to determine the constant load PO that would elicit the  $\dot{V}O_2$  response at RCP,  $m[\text{HHb}]_{\text{BP}}$  and  $i\text{EMG}_{\text{BP}}$ , a novel correction strategy was applied (Caen et al. 2020). In specific, the corrected PO for constant load exercise was calculated by combining a simple MRT correction with an additional correction that accounted for the loss of mechanical efficiency at higher intensities, given by the following formula:

$$\text{PO} = (\text{PO}_{\text{ramp}} - \text{MRT}) - \left( \frac{(\text{PO}_{\text{ramp}} - \text{GET}) \times (14.2 - s_{2\text{-ramp}})}{14.2} \right),$$

where  $\text{PO}_{\text{ramp}}$  is the ramp-identified PO for each threshold, GET is the uncorrected PO, and the constant of  $14.2 \text{ mL min}^{-1} \text{W}^{-1}$  corresponds to the estimated  $\dot{V}O_2$  gain during constant load exercise above GET (see Caen et al. (2020) for a detailed description).

### CP tests

CP and  $W'$  were estimated using a ‘best individual fit’ approach, meaning that the regression model associated with the lowest total error in estimating CP and  $W'$  (i.e., sum of the standard errors) was selected on an individual basis. Three mathematical models were included in the analysis: the two-parameter hyperbolic model (time vs. PO), the linear total work model (total work vs. PO) and the linear 1/time model (PO vs. 1/time). Models were only eligible if they met the following criteria: (1) a minimum of three trials were included, (2) time to exhaustion of the trials was proportionally distributed between 2 and 20 min with at least one trial that was longer than 12 min (Mattioni Maturana et al. 2018), (3) the end-exercise  $\dot{V}O_2$  of each trial exceeded 95%  $\dot{V}O_{2\text{peak}}$ , and (4) the standard error of estimate for CP and  $W'$ , expressed relative as a coefficient of variation (CV%),

was  $< 5\%$  and  $< 10\%$ , respectively. The  $\dot{V}O_2$  at CP was calculated as the average value during the final 5 min of the first 30-min constant load test.

### MLSS trials

MLSS was defined as the highest PO that could be sustained for 30 min with a steady-state  $\text{La}^-$  response, defined as an increase  $\leq 1 \text{ mM}$  between the 10th and 30th min (Heck et al. 1985). If  $[\text{La}^-]$  was stable, the subsequent test was performed at CP + 10 W. If  $[\text{La}^-]$  increased by  $> 1 \text{ mM}$ , the following test was performed at CP – 10 W. This procedure was repeated until the highest PO associated with a steady-state blood  $\text{La}^-$  response was determined. In the event of edge cases, the corresponding  $\text{La}^-$  curves were evaluated by three experts and it was checked whether the  $\text{La}^-$  response could be indicative for the achievement of a delayed steady state (Jones et al. 2019). The  $\dot{V}O_2$  at MLSS was defined as the average value during the final 5 min of exercise.

### Statistical analysis

Statistical analyses were performed separately for group I ( $n = 10$ ) and II ( $n = 11$ ) using SPSS Statistics 24 (IBM Corp, Armonk, NY). Descriptive data are presented as mean  $\pm$  SD. Statistical significance was accepted at  $P < 0.05$ .

A two-way repeated measures ANOVA [time (2)  $\times$  CP model (3)] was executed to compare CP estimates between mathematical models. Furthermore, paired-samples  $t$ -tests were used to compare parameters of performance and cardiorespiratory fitness between the two ramp incremental tests and to compare  $W'$  values between T1 and T2.

In both studies, a two-way repeated measures ANOVA [time (2)  $\times$  thresholds (5)] was conducted to compare the thresholds over time. This analysis was performed two times, once for the thresholds expressed as  $\dot{V}O_2$  and once for the thresholds expressed as PO. Greenhouse–Geisser corrections were applied when Mauchly’s test of sphericity indicated that the assumption of sphericity had been violated. Post hoc tests with LSD adjustment were used for multiple comparisons.

Test–retest reliability of the thresholds in study 1 was evaluated by calculating the intraclass correlation coefficients (ICC) and its 95% confidence intervals ( $\text{CI}_{95\%}$ ) using a single-measurement, absolute-agreement, two-way mixed-effects model (Koo and Li 2016). For each threshold, the typical error of measurement (TEM) was calculated as the SD of the difference score divided by  $\sqrt{2}$ , and then expressed as a coefficient of variation (CV%) after log-transformation of the variables (Hopkins 2000). The minimum difference (MD) to be considered a real change in each threshold on the

95% confidence level was calculated by multiplying TEM with  $1.96 \times \sqrt{2}$  (Weir 2005). The Pearson correlation coefficient ( $r$ ) was used to detect significant associations between the thresholds in both studies, and between their temporal changes ( $\Delta$ ) from T1 to T2 in study 2.

## Results

### Study 1: 4-week test–retest reliability study ( $n = 10$ women)

Table 1 gives an overview of the performance and cardiorespiratory parameters derived from both ramp tests. (Sub)maximal performance parameters did not improve from T1 to T2 ( $P > 0.05$ ).

CV% associated with CP and  $W'$  were, respectively,  $0.7 \pm 0.8\%$  and  $3.5 \pm 2.3\%$  at T1, and  $1.0 \pm 0.8\%$  and  $4.9 \pm 2.9\%$  at T2. For all tests, best model fits were derived from the hyperbolic model in nine cases and from the linear 1/time model in eleven cases. There were no significant differences between CP estimates derived from the different mathematical models ( $P = 0.249$ ).  $W'$  did not change from T1 ( $14.6 \pm 3.7$  kJ) to T2 ( $15.2 \pm 6.1$  kJ) ( $P = 0.672$ ). All participants were able to complete 30 min of cycling at CP, except for one at T2 (time to exhaustion = 23.6 min). In 40% of the tests (8/20), cycling at CP was not associated with a  $La^-$  steady-state response (i.e.,  $+2.24 \pm 0.6$  mM between the 10th and 30th min), in 50% of the tests (10/20), CP was equal to MLSS, while in the remaining 10% of cases (2/20), MLSS was higher than CP. MLSS occurred at blood [ $La^-$ ] of  $6.2 \pm 1.3$  mM and was associated with an average increase of  $0.5 \pm 0.3$  mM during the final 20 min of exercise.

Table 2 compares CP, MLSS, RCP,  $m[HHb]_{BP}$  and  $iEMG_{BP}$  when expressed as  $\dot{V}O_2$  and in PO. There was

**Table 2** Comparison of  $\dot{V}O_2$  and PO values corresponding to CP, MLSS, RCP,  $m[HHb]_{BP}$  and  $iEMG_{BP}$  at both timepoints (T1 and T2) in study 1 ( $n = 10$ )

|   | Grand mean  | T1          | T2          | $\Delta$    |
|---|-------------|-------------|-------------|-------------|
| Thresholds expressed as $\dot{V}O_2$ (L min <sup>-1</sup> ) |             |             |             |             |
| Grand mean  |             | 2.35 (0.14) | 2.41 (0.14) | 0.06 (0.04) |
| (1) CP  | 2.43 (0.15) | 2.40 ± 0.46 | 2.45 ± 0.52 | 0.05 ± 0.16 |
| (2) MLSS  | 2.35 (0.15) | 2.34 ± 0.49 | 2.37 ± 0.45 | 0.03 ± 0.12 |
| (3) RCP   | 2.33 (0.16) | 2.30 ± 0.54 | 2.37 ± 0.48 | 0.07 ± 0.18 |
| (4)   | 2.36 (0.10) | 2.31 ± 0.35 | 2.40 ± 0.33 | 0.09 ± 0.28 |
| $m[HHb]_{BP}$   |             |             |             |             |
| (5) $iEMG_{BP}$   | 2.44 (0.15) | 2.43 ± 0.48 | 2.46 ± 0.51 | 0.03 ± 0.27 |
| Thresholds expressed as PO (W)                              |             |             |             |             |
| Grand mean  |             | 173 (12)    | 175 (11)    | 2 (2)       |
| (1) CP  | 177 (13)    | 174 ± 39    | 180 ± 42    | 5 ± 9       |
| (2) MLSS  | 173 (12)    | 171 ± 41    | 175 ± 37    | 3 ± 8       |
| (3) RCP   | 170 (12)    | 170 ± 42    | 171 ± 36    | 1 ± 12      |
| (4)   | 172 (8)     | 171 ± 28    | 172 ± 26    | 2 ± 16      |
| $m[HHb]_{BP}$   |             |             |             |             |
| (5) $iEMG_{BP}$   | 177 (12)    | 179 ± 37    | 176 ± 38    | 3 ± 15      |

Data are presented as mean ± SD. Grand means represent the estimated marginal means (standard error) from the RM ANOVA. No significant differences are reported

$\dot{V}O_2$  oxygen uptake, PO power output, CP critical power, MLSS maximal lactate steady state, RCP respiratory compensation point,  $m[HHb]_{BP}$  breakpoint in deoxygenated hemoglobin,  $iEMG_{BP}$  breakpoint in electromyography

no significant time × thresholds interaction effect ( $\dot{V}O_2$ :  $P = 0.803$  and  $\eta_p^2 = 0.025$ , PO:  $P = 0.547$  and  $\eta_p^2 = 0.08$ ). Furthermore, we did not find a main effect for time ( $\dot{V}O_2$ :  $P = 0.233$  and  $\eta_p^2 = 0.154$ , PO:  $P = 0.546$  and  $\eta_p^2 = 0.042$ ), nor for the thresholds ( $\dot{V}O_2$ :  $P = 0.143$  and  $\eta_p^2 = 0.169$ , PO:  $P = 0.281$  and  $\eta_p^2 = 0.132$ ), indicating that the thresholds did not significantly change between T1 and T2, and that there were no significant differences between

**Table 1** Peak performance and cardiorespiratory parameters derived from the ramp incremental exercise tests at both timepoints (T1 and T2) in study 1 ( $n = 10$ ) and study 2 ( $n = 11$ )

|   | Study 1 ( $n = 10$ ) |              | Study 2 ( $n = 11$ ) |              |
|---|----------------------|--------------|----------------------|--------------|
|   | T1                   | T2           | T1                   | T2           |
| $PO_{peak}$ (W)   | 287 ± 47             | 291 ± 45     | 380 ± 40             | 395 ± 50*    |
| $\dot{V}O_{2peak}$ (mL min <sup>-1</sup> )                  | 2.62 ± 0.57          | 2.67 ± 0.50  | 3.69 ± 0.51          | 3.90 ± 0.47* |
| $\dot{V}O_{2peak}$ (mL min <sup>-1</sup> kg <sup>-1</sup> ) | 43.2 ± 7.3           | 44.4 ± 6.8   | 47.7 ± 5.9           | 50.8 ± 5.6*  |
| $HR_{peak}$ (bpm)   | 185 ± 12             | 182 ± 11*    | 189 ± 12             | 186 ± 11     |
| $RER_{peak}$  | 1.31 ± 0.09          | 1.23 ± 0.06* | 1.39 ± 0.08          | 1.36 ± 0.08  |
| $\dot{V}E_{peak}$   | 118 ± 20             | 113 ± 19     | 160 ± 23             | 162 ± 28     |
| GET (W)   | 124 ± 34             | 126 ± 27     | 175 ± 45             | 180 ± 38     |
| MRT (s)   | 42 ± 8               | 41 ± 7       | 52 ± 6               | 47 ± 12      |

PO power output,  $\dot{V}O_2$  pulmonary oxygen uptake, HR heart rate, RER respiratory exchange ratio,  $\dot{V}E$  pulmonary ventilation, GET gas exchange thresholds, MRT mean response time

\*denotes significant differences between T1 and T2 ( $P < 0.05$ )

them. Assessment of the 4-week test–retest reliability of the thresholds is presented in Table 3. All thresholds were strongly correlated at T1 and T2 when they were expressed as  $\dot{V}O_2$  ( $r=0.75$  to  $0.98$ ) and  $PO$  ( $r=0.86$  to  $0.99$ ) (Fig. 1A).

### Study 2: 12-week intervention study ( $n=11$ men)

Table 1 gives an overview of the performance and cardiorespiratory parameters derived from both ramp tests.

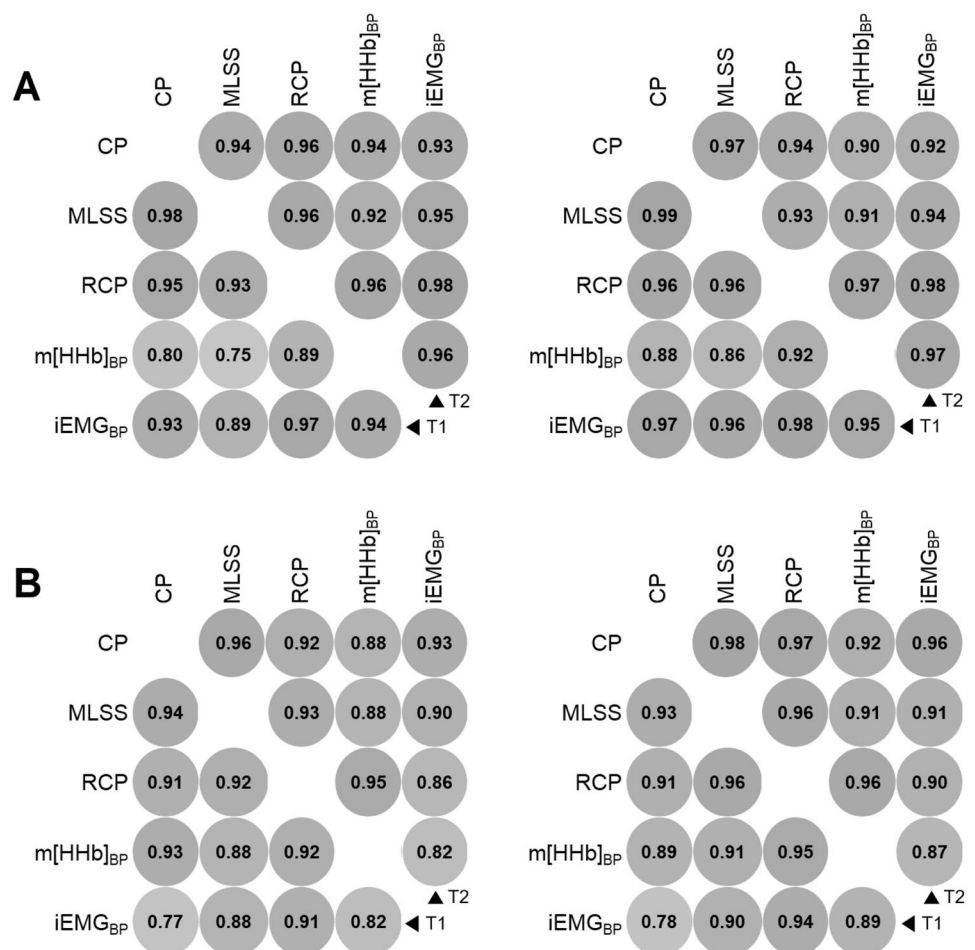
**Table 3** Parameters evaluating 4-week test–retest reliability of CP, MLSS, RCP, m[HHb]<sub>BP</sub> and iEMG<sub>BP</sub> in study 1 ( $n=10$ )

|                             | CP                | MLSS              | RCP               | m[HHb] <sub>BP</sub> | iEMG <sub>BP</sub> |
|-----------------------------|-------------------|-------------------|-------------------|----------------------|--------------------|
| $\dot{V}O_2$                |                   |                   |                   |                      |                    |
| TEM (mL min <sup>-1</sup> ) | 111 [76, 202]     | 88 [60, 160]      | 125 [86, 228]     | 195 [134, 357]       | 189 [130, 346]     |
| CV (%)                      | 4.5 [3.1, 8.5]    | 4.0 [2.7, 7.4]    | 5.7 [3.9, 10.6]   | 9.6 [6.5, 18.3]      | 8.8 [6.0, 16.7]    |
| MD (mL min <sup>-1</sup> )  | 308               | 244               | 346               | 541                  | 524                |
| ICC                         | 0.95 [0.82, 0.99] | 0.97 [0.88, 0.99] | 0.94 [0.77, 0.98] | 0.67 [0.15, 0.91]    | 0.87 [0.55, 0.97]  |
| $PO$ (W)                    |                   |                   |                   |                      |                    |
| TEM (W)                     | 7 [5, 12]         | 6 [4, 11]         | 10 [7, 18]        | 18 [12, 33]          | 15 [10, 26]        |
| CV (%)                      | 3.6 [2.5, 6.8]    | 3.5 [2.4, 6.5]    | 4.3 [2.9, 7.9]    | 9.0 [6.1, 17.0]      | 7.0 [4.8, 13.1]    |
| MD (W)                      | 19                | 17                | 28                | 50                   | 42                 |
| ICC                         | 0.97 [0.87, 0.99] | 0.98 [0.91, 0.99] | 0.96 [0.85, 0.99] | 0.58 [-0.06, 0.88]   | 0.89 [0.62, 0.97]  |

Data are presented as mean [CI<sub>95%</sub>]

TEM typical error of measurement, CV% coefficient of variation, ICC intraclass correlation coefficient, MD minimum difference

**Fig. 1** Correlation matrix showing the associations between CP, MLSS, RCP, m[HHb]<sub>BP</sub> and iEMG<sub>BP</sub> at both timepoints (T1 and T2) in study 1 ( $n=10$ ) (A) and study 2 ( $n=11$ ) (B). Threshold correlations were examined for their associated  $\dot{V}O_2$  (left matrices) and their corresponding  $PO$  (right matrices). Significant correlations ( $P<0.05$ ) are in bold



After 12 weeks,  $PO_{\text{peak}}$  ( $3.5 \pm 5.2\%$ ,  $P=0.042$ ) and  $\dot{V}O_{2\text{peak}}$  ( $6.9 \pm 6.2\%$ ,  $P=0.005$ ) were significantly increased.

The CV% associated with CP and W' were, respectively,  $0.5 \pm 0.4\%$  and  $3.6 \pm 2.2\%$  at T1, and  $1.1 \pm 0.7\%$  and  $6.4 \pm 4.0\%$  at T2. For all tests, best model fits were derived from the hyperbolic model in 12 cases and from the linear 1/time model in 10 cases. There were no significant differences between CP estimates derived from the different mathematical models ( $P=0.314$ ). W' was decreased at T2 ( $20.8 \pm 5.2$  kJ) compared to T1 ( $22.9 \pm 4.1$  kJ) ( $P=0.023$ ). One participant was not able to complete 30 min of CP exercise at T1 (time to exhaustion = 20.7 min), while this was the case for three participants at T2 (time to exhaustion = 17.3 min, 20.2 min and 21.5 min). In 50% of the tests (11/22), cycling at CP was not associated with a  $La^-$  steady-state response (i.e.,  $+2.67 \pm 1.49$  mM between the 10th and 30th min), whereas in 45% of the tests (10/22), CP was equal to MLSS. In only one case, MLSS was found to be higher than CP. MLSS occurred at blood  $[La^-]$  of  $5.8 \pm 1.9$  mM and was associated with an average increase of  $0.45 \pm 0.45$  mM during the final 20 min of exercise.

Table 4 compares CP, MLSS, RCP,  $m[HHb]_{\text{BP}}$  and  $iEMG_{\text{BP}}$  when expressed as  $\dot{V}O_2$  and in PO. There was no significant time  $\times$  thresholds interaction effect ( $\dot{V}O_2$ :  $P=0.487$  and  $\eta_p^2=0.071$ , PO:  $P=0.537$  and  $\eta_p^2=0.063$ ). However, a significant main effect for time showed that the  $\dot{V}O_2$  and PO values associated with the thresholds were higher at T2 compared to T1 ( $\dot{V}O_2$ :  $P=0.001$  and  $\eta_p^2=0.674$ , PO:  $P=0.026$  and  $\eta_p^2=0.405$ ). In addition, we found a

significant main effect for the thresholds ( $\dot{V}O_2$ :  $P<0.001$  and  $\eta_p^2=0.427$ , PO:  $P=0.024$  and  $\eta_p^2=0.239$ ), revealing the presence of significant differences between them. In specific, when expressed as  $\dot{V}O_2$  ( $L \cdot \text{min}^{-1}$ ),  $m[HHb]_{\text{BP}}$  was higher than CP ( $P=0.002$ ,  $CI_{95\%}=84$  to  $266 \text{ mL} \cdot \text{min}^{-1}$ ), MLSS ( $P<0.001$ ,  $CI_{95\%}=140$  to  $326 \text{ mL} \cdot \text{min}^{-1}$ ) and RCP ( $P=0.026$ ,  $CI_{95\%}=18$  to  $218 \text{ mL} \cdot \text{min}^{-1}$ ). Furthermore,  $iEMG_{\text{BP}}$  was higher than MLSS ( $P=0.006$ ,  $CI_{95\%}=52$  to  $236 \text{ mL} \cdot \text{min}^{-1}$ ). When the thresholds were expressed as PO (W), CP was higher than MLSS ( $P=0.031$ ,  $CI_{95\%}=1$  to  $15$  W), and  $m[HHb]_{\text{BP}}$  was higher than MLSS ( $P=0.021$ ,  $CI_{95\%}=2$  to  $18$  W) and RCP ( $P=0.019$ ,  $CI_{95\%}=2$  to  $16$  W). All thresholds were strongly correlated at T1 and T2 when they were associated by their  $\dot{V}O_2$  ( $r=0.77$  to  $0.96$ ) and PO ( $r=0.78$  to  $0.98$ ) (Fig. 1B). Correlations between the  $\dot{V}O_2$  and PO changes in the thresholds (i.e.,  $\Delta$  values from T1 to T2) are shown in Fig. 2.

## Discussion

This study is the first to make a comprehensive evaluation of the interrelationship and reproducibility of CP, MLSS, RCP,  $m[HHb]_{\text{BP}}$  and  $iEMG_{\text{BP}}$ . Thresholds were examined by means of two longitudinal studies: a 4-week test–retest reliability study ( $n=10$ ) and a 12-week intervention study ( $n=11$ ). The main goal was to investigate whether these thresholds could be used interchangeably in

**Table 4** Comparison of  $\dot{V}O_2$  and PO values corresponding to CP, MLSS, RCP,  $m[HHb]_{\text{BP}}$  and  $iEMG_{\text{BP}}$  at both timepoints (T1 and T2) in study 2 ( $n=11$ )

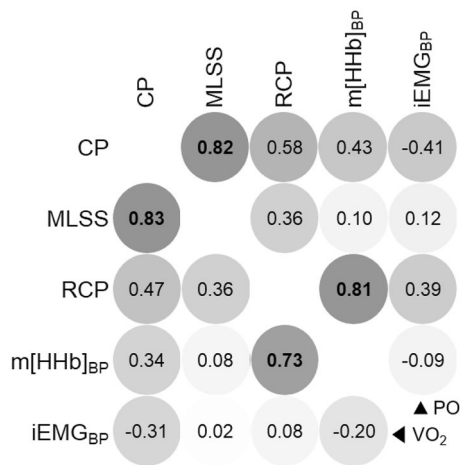
|  | Grand mean                   | T1              | T2              | $\Delta$        |
|--|------------------------------|-----------------|-----------------|-----------------|
| Thresholds expressed as $\dot{V}O_2$ ( $L \cdot \text{min}^{-1}$ ) |                              |                 |                 |                 |
| Grand mean   |                              | 3.35 (0.12)     | 3.53 (0.12)*    | 0.18 (0.04)     |
| (1) CP   | 3.38 (0.10) <sup>4</sup>     | 3.31 $\pm$ 0.33 | 3.46 $\pm$ 0.33 | 0.15 $\pm$ 0.18 |
| (2) MLSS   | 3.33 (0.10) <sup>4,5</sup>   | 3.24 $\pm$ 0.34 | 3.41 $\pm$ 0.36 | 0.18 $\pm$ 0.16 |
| (3) RCP  | 3.44 (0.14) <sup>4</sup>     | 3.34 $\pm$ 0.49 | 3.55 $\pm$ 0.48 | 0.21 $\pm$ 0.22 |
| (4) $m[HHb]_{\text{BP}}$   | 3.56 (0.12) <sup>1,2,3</sup> | 3.43 $\pm$ 0.45 | 3.69 $\pm$ 0.40 | 0.25 $\pm$ 0.27 |
| (5) $iEMG_{\text{BP}}$   | 3.47 (0.13) <sup>2</sup>     | 3.41 $\pm$ 0.46 | 3.53 $\pm$ 0.45 | 0.11 $\pm$ 0.28 |
| Thresholds expressed as PO (W)                                     |                              |                 |                 |                 |
| Grand mean   |                              | 243 (12)        | 254 (12)*       | 11 (4)          |
| (1) CP   | 252 (11) <sup>2</sup>        | 245 $\pm$ 37    | 259 $\pm$ 40    | 13 $\pm$ 18     |
| (2) MLSS   | 244 (12) <sup>1,4</sup>      | 239 $\pm$ 38    | 249 $\pm$ 44    | 10 $\pm$ 16     |
| (3) RCP  | 245 (12) <sup>4</sup>        | 239 $\pm$ 42    | 251 $\pm$ 40    | 12 $\pm$ 18     |
| (4) $m[HHb]_{\text{BP}}$   | 254 (11) <sup>2,3</sup>      | 246 $\pm$ 41    | 261 $\pm$ 36    | 15 $\pm$ 23     |
| (5) $iEMG_{\text{BP}}$   | 247 (11)                     | 245 $\pm$ 41    | 250 $\pm$ 39    | 5 $\pm$ 25      |

Data are presented as mean  $\pm$  SD. Grand means represent the estimated marginal means (standard error) from the RM ANOVA and are used to denote significant differences. Threshold numbers in superscript <sup>(1–5)</sup> specify the thresholds from which they differ ( $P<0.05$ )

$\dot{V}O_2$  oxygen uptake, PO power output, CP critical power, MLSS maximal lactate steady state, RCP respiratory compensation point,  $m[HHb]_{\text{BP}}$  breakpoint in deoxygenated hemoglobin,  $iEMG_{\text{BP}}$  breakpoint in electromyography

\*denotes significant differences between T1 and T2 ( $P<0.05$ )





**Fig. 2** Correlation matrix showing the associations between the changes in the thresholds over time in study 2 ( $n = 11$ ). Threshold changes over 8 weeks were examined for their associated  $\dot{V}O_2$  and their corresponding PO. Significant correlations ( $P < 0.05$ ) are in bold

practice to estimate the MMSS intensity. For this purpose, thresholds were compared based on their associated  $\dot{V}O_2$  and their equivalent PO for constant load exercise. In study 1, we demonstrated that the whole-body thresholds (i.e., CP, MLSS and RCP) showed excellent test–retest reliability across a 4-week timespan (as reflected by TEM, CV%, MD and ICC), whereas the test–retest reliability for the local thresholds (i.e., iEMG<sub>BP</sub> and m[HHb]<sub>BP</sub>) was considerably lower. In addition, we found that all thresholds were strongly associated and that they did not differ, regardless of their expression method (i.e., as  $\dot{V}O_2$  or PO). In study 2, we found significant differences in the  $\dot{V}O_2$  and PO values between the whole-body and local thresholds. In addition, apart from a significant correlation between the temporal changes ( $\Delta$ ) in CP and MLSS, and between  $\Delta$ RCP and  $\Delta$ m[HHb]<sub>BP</sub>, none of the other  $\Delta$ -values were significantly associated between thresholds. Together, although the determination of CP, MLSS, RCP, m[HHb]<sub>BP</sub> and iEMG<sub>BP</sub> may lead to similar  $\dot{V}O_2$  and PO values, the present study results demonstrate that whole-body and local thresholds should not be used interchangeably in practice.

### Whole-body thresholds

Although CP and MLSS are both considered as reference methods to determine the MMSS intensity, these thresholds are two different concepts with distinct methodologies. A recent meta-analysis that combined the results of five comparative studies reported that CP occurred  $\sim 11\%$  higher than the PO at MLSS (Galán-Rioja et al. 2020). In the present work, there was no significant difference between CP and MLSS in study 1 ( $\Delta = 4$  W, Table 2), while there was in study 2 ( $\Delta = 8$  W, Table 4). These differences ( $\sim 6$  W or 3%

on average) are considerably smaller than what has been found in previous studies (Galán-Rioja et al. 2020). The only other study that does not report significant differences between CP and MLSS is the one from Keir et al. (2015b) ( $\Delta = 2$  W). However, in that study, CP was modeled by fitting a three-parameter hyperbolic model, which typically yields lower CP estimates compared to the traditional two-parameter models (Bull et al. 2000; Bergstrom et al. 2014). Thus, this study is the first to find a very close correspondence between CP and MLSS when CP is determined from the original two-parameter models.

In fact, much of the discrepancy between CP and MLSS reported in the literature can be attributed to the methodology of their determination and therefore, differences between both thresholds should always be interpreted in relation to the precision with which they were determined. In case of CP, errors can arise from the accuracy of the mathematical modeling, which is highly dependent on the duration and the number of predictive trials that are used (Muniz-Pumares et al. 2019). In this study, we used very strict criteria for determining CP and  $W'$ , which led to excellent model fittings with an average CV% of 0.8% and 4.6% for CP and  $W'$ , respectively. With respect to MLSS, methodological issues are particularly related to the precision and interpretation of the blood  $La^-$  measurements, and to the size of the PO increments or decrements between successive tests. As a direct consequence of this latter, the PO at MLSS will actually be an underestimate of the true MLSS (Jones et al. 2019). Still, using small steps (i.e., 10 W in this study), this underestimation can be limited to a few watts at maximum (i.e., 0 to 9 W). Overall, we observed that CP and MLSS occurred in very close vicinity in almost all participants, with only 3 out of 42 comparisons where the discrepancy between both was  $> 10$  W. Considering the methodological background, the presence of tiny differences between CP and MLSS, whether they are statistically significant or not, should be expected.

Interestingly, we found that exercise at CP did not evoke a  $La^-$  steady-state response in 45% of all tests (i.e., a blood  $[La^-]$  increase of  $2.38 \pm 1.06$  mM between 10 and 30th min) and was associated with high end-values for  $\dot{V}O_2$  ( $92 \pm 5\% \dot{V}O_{2peak}$ ), HR ( $96 \pm 4\% HR_{peak}$ ) and RPE ( $18 \pm 2$ ). These responses are very similar to the findings of Brickley et al. (2002), but do not necessarily mean that CP is invalid to estimate the MMSS intensity. This could be explained by the idea that it is impossible to exercise exactly “at” the threshold. Jones et al. (2019) reasoned that if CP represents the actual heavy-to-severe exercise boundary, cycling at this exact CP would in reality lead to a 50%-chance that someone is exercising just below CP (i.e., heavy domain) or just above CP (i.e., severe domain), in each case leading to distinct physiological responses. The fact that the physiological response “at” CP did not seem to indicate a steady-state

response in all participants could thus be a direct consequence of some participants who cycled above their actual MMSS (i.e., within the severe domain).

Both CP and MLSS demonstrated excellent 4-week test–retest reliability (Table 3), did not differ from each other in terms of  $\dot{V}O_2$  (Tables 2 and 4), and were among the only thresholds that showed strong associations ( $r=0.82$ – $0.83$ ) between their changes from T1 to T2 (Fig. 2). All these findings support the notion that CP and MLSS can be used to estimate the MMSS intensity, taking into account a certain “grey zone” of natural and methodological variation in which the thresholds may be located.

In agreement with previous studies, we found that the  $\dot{V}O_2$  response at RCP did not differ from CP or MLSS (Tables 2 and 4) (Keir et al. 2015b; Mattioni Maturana et al. 2016; Iannetta et al. 2018; Inglis et al. 2020; Caen et al. 2020, 2021). Furthermore, translating RCP into its appropriate PO for constant load exercise revealed no significant differences with CP or MLSS as well. Again, this confirms that previously reported differences between RCP and CP/MLSS were due to issues with accounting for the gap in the  $\dot{V}O_2$ /PO relationship between ramp and constant load exercise (Keir et al. 2018a; Caen et al. 2020). Despite the fact that the translation of RCP into its equivalent PO includes several methodological steps (see “Methods” section), thereby increasing the risk for accumulating small calculation errors, RCP showed excellent reliability (Table 2). Therefore, our results reinforce the idea of utilizing RCP as an alternative to CP/MLSS when the MMSS intensity needs to be estimated from a single incremental test.

### Local thresholds

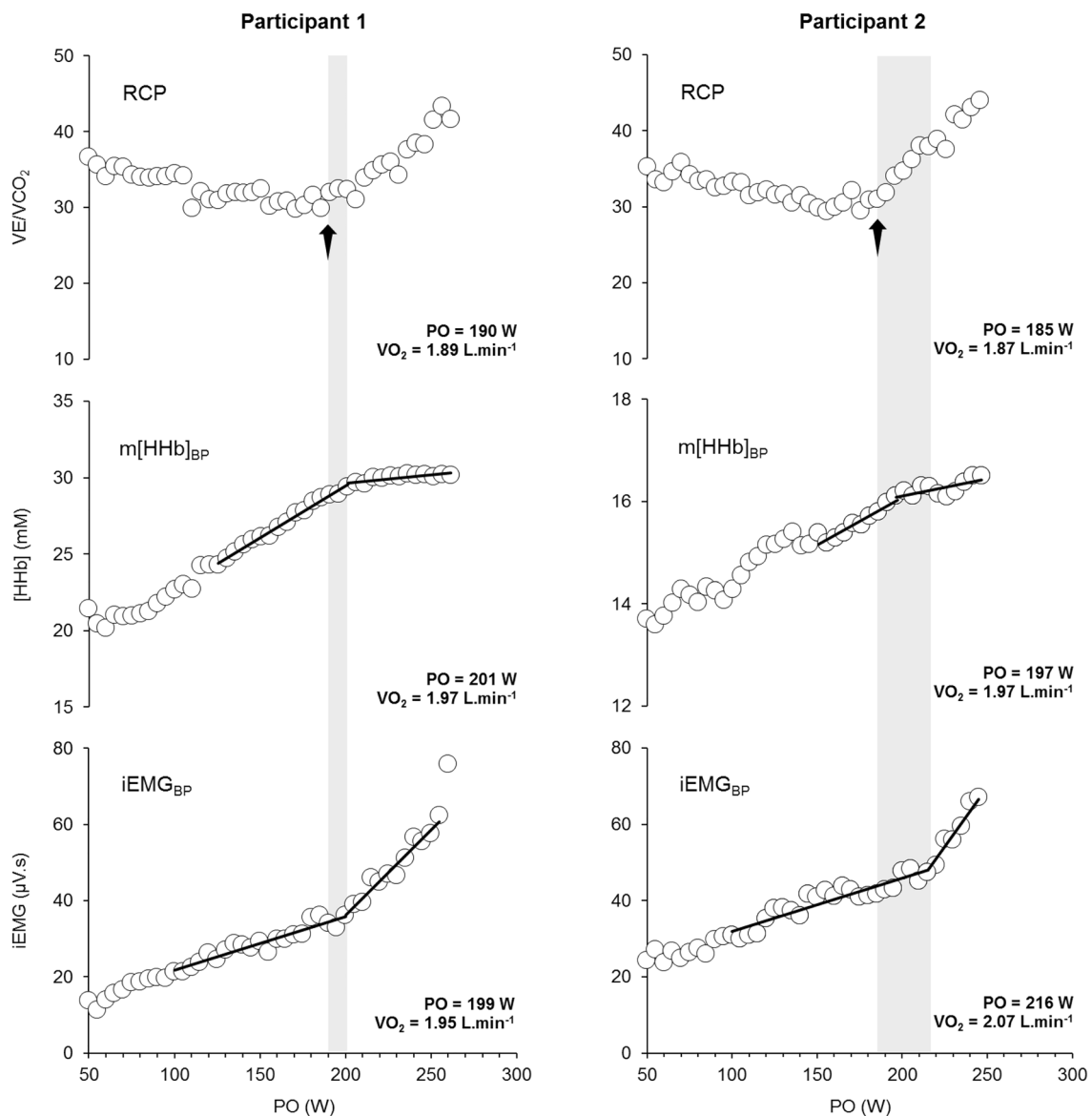
The occurrence of CP, MLSS, RCP,  $m[HHb]_{BP}$  and  $iEMG_{BP}$  within a very narrow range of exercise intensities ( $\sim 88$  to  $93\% \dot{V}O_{2peak}$ ) and the fact that they were all strongly correlated ( $r=0.75$ – $0.99$ ) supports the idea that these thresholds are collectively linked to the physiological events that occur during the transition from heavy-to-severe exercise. In this light, Boone et al. (2016) has previously presented a framework describing the mechanistic link between them. Still, it should be emphasized that a physiological relation alone does not mean that the different thresholds can be used interchangeably in practice.

In contrast to the whole-body thresholds, the study outcomes with regard to  $m[HHb]_{BP}$  and  $iEMG_{BP}$  are less consistent. While we did not find differences between whole-body and local exercise thresholds in study 1 (Table 2),  $m[HHb]_{BP}$  and  $iEMG_{BP}$  showed significant deviations in study 2 (Table 4). In spite of statistical significance, it could be questioned whether the magnitude of these differences ( $\Delta = 118$ – $233$  mL  $\text{min}^{-1}$

or 9–10 W) is physiologically meaningful considering the normal variation expected from gas exchange measurements (i.e.,  $\sim 140$  mL  $\text{min}^{-1}$  during steady-state exercise according to Keir et al. (2014)) and taking into account the reliability of these local thresholds. Indeed, TEM for  $m[HHb]_{BP}$  and  $iEMG_{BP}$  was markedly higher (i.e.,  $\sim 190$  mL  $\text{min}^{-1}$  or 15–18 W) compared to the whole-body thresholds and could, according to the  $CI_{95\%}$ , reach values greater than 350 mL  $\text{min}^{-1}$  or 30 W (Table 3). In addition, we found a wide range of ICC ( $CI_{95\%} = -0.06$  to 0.97) for both thresholds and the calculated MD indicated that changes in  $m[HHb]_{BP}$  and  $iEMG_{BP}$  would need to exceed 500 mL  $\cdot$   $\text{min}^{-1}$  or 40–50 W before they could be considered “real” with 95% confidence. These results seriously challenge the usability of  $m[HHb]_{BP}$  and  $iEMG_{BP}$  for training practice, regardless of their coincidence with the other thresholds. In this context, we should also keep in mind that local thresholds such as  $m[HHb]_{BP}$  and  $iEMG_{BP}$  are based on the premise that the selected muscle regions (i.e., M. Vastus Lateralis in this study) are representative of the overall recruited muscle response. However, it is known that there is spatial heterogeneity in the activation and oxygenation pattern within the exercising muscles (Koga et al. 2007, 2011), so that the identification of local thresholds may depend on the specific area of muscle fibers that are measured.

Even if some of the “significant” differences between CP/MLSS/RCP and  $m[HHb]_{BP}/iEMG_{BP}$  may be attributed to measurement errors or typical variation in these measurements, a closer look at our data suggest that they do not necessarily represent just random variation. To illustrate this, we have presented the threshold determination of two female participants in Fig. 3. As shown, there is relatively little noise on the data and kinetics are clear, ensuring an accurate determination of RCP,  $m[HHb]_{BP}$  and  $iEMG_{BP}$ . We see that the range of intensities within which the thresholds occur varies considerably between these two participants. While in participant 1, the thresholds occurred more or less simultaneously, a more sequential and spread threshold occurrence was visible in participant 2. These examples may suggest that the interrelationship of thresholds might be dependent on individual characteristics as well.

In study 2, we did not find a significant interaction between the thresholds and their evolution over time, which would mean that their interrelationship remains stable. However, we must note that non-significant interaction effects are often a consequence of low statistical power and do not prove that there is no effect modification. In a previous study, we observed that none of the changes in the thresholds was correlated and that there was a high level of variability in the way thresholds had



**Fig. 3** Occurrence of RCP,  $m[\text{HHb}]_{\text{BP}}$  and  $i\text{EMG}_{\text{BP}}$  in two participants demonstrating a high level of interindividual variability. In participant 1, the thresholds occur nearly simultaneously ( $\sim 20$  s) and their associated  $\dot{V}\text{O}_2$  response closely corresponds to MLSS

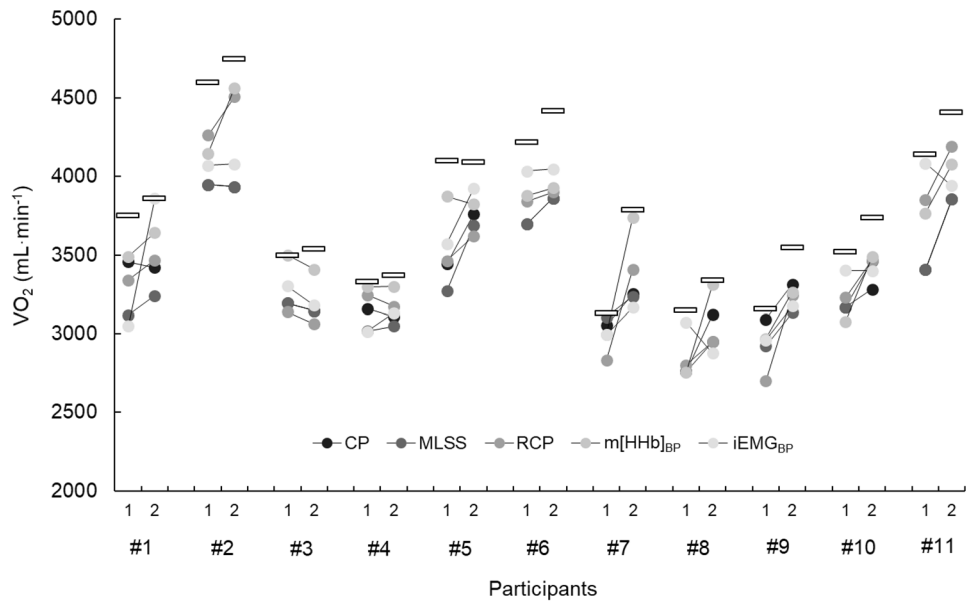
(mean  $\Delta = 20$  mL  $\text{min}^{-1}$ ). In participant 2, the thresholds occur in a sequence over a longer time window ( $\sim 60$  s) and their associated  $\dot{V}\text{O}_2$  response does not show a close correspondence to MLSS (mean  $\Delta = 180$  mL  $\text{min}^{-1}$ )

evolved over time (Caen et al. 2018). In contrast, Inglis et al. (2020) reported high correlations between the  $\dot{V}\text{O}_2$  changes in MLSS, RCP and  $m[\text{HHb}]_{\text{BP}}$ . Figure 2 shows that the temporal changes in the thresholds were weakly associated, suggesting an inconsistent interrelationship of the thresholds. Taken together, our results demonstrate a significant level of both inter- and intrasubject variability in the thresholds, in particular with reference to  $m[\text{HHb}]_{\text{BP}}$  and  $i\text{EMG}_{\text{BP}}$ . This high variability in the temporal changes of the thresholds is clearly visible in Fig. 4.

## Conclusion

Since the early introduction of exercise thresholds, many researchers have studied the equivalence of different threshold concepts and examined their association with the MMSS intensity. Although it is a repetitive finding in the literature that these thresholds are well correlated and occur within a narrow range of exercise intensities, their equivalence and interchangeable use remains controversial. In spite of the fact that CP, MLSS, RCP,  $m[\text{HHb}]_{\text{BP}}$  and  $i\text{EMG}_{\text{BP}}$  were well correlated and occurred in each other's close proximity, a finding that strengthens the

**Fig. 4** Overview of the individual changes in CP, MLSS, RCP,  $m[HHb]_{BP}$  and  $iEMG_{BP}$  from timepoint 1 to timepoint 2 in study 2 ( $n = 11$ ). Threshold values are expressed as  $\dot{V}O_2$  ( $mL \text{ min}^{-1}$ ). Horizontal stripes denote  $\dot{V}O_{2peak}$  as derived from the ramp test. Note that the thresholds evolved very similarly in some participants, whereas their observed changes were much more independent in others

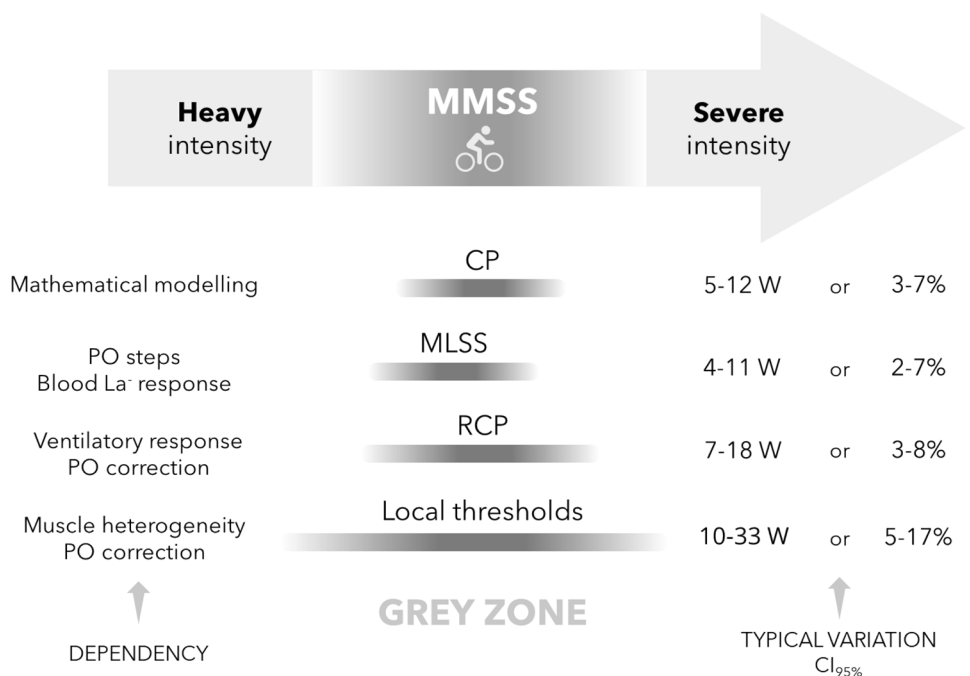


conviction that these thresholds are expressions of linked physiological phenomena, our study results do not provide a cogent argument for their equivalence. Still, the rather limited resolution with which we are currently able to determine thresholds together with the low reproducibility of some of them limits our ability to discover the exact temporal sequence of physiological events inside the body when crossing the heavy-to-severe exercise boundary and therefore, make it dangerous to draw definite conclusions about their true equivalence. By any means, we should always be aware of the presence of a certain “grey zone”

surrounding the thresholds, accounting for both natural and methodological variation. The fact that we observed that all exercise thresholds occurred within a narrow range of exercise intensities, but not simultaneously per se, corresponds to this idea. Based on the present study results, Fig. 5 presents a conceptual framework visualizing the position of the investigated thresholds within a bandwidth of exercise intensities surrounding the actual MMSS intensity.

In summary, we demonstrated a close and consistent coherence between CP, MLSS and RCP, provided that these

**Fig. 5** Visualization of critical power (CP), maximal lactate steady state (MLSS), respiratory compensation point (RCP) and local thresholds within a presumed “grey zone” of exercise intensities surrounding the actual maximal metabolic steady-state (MMSS) intensity. The exact bandwidth of intensities for each threshold is defined according to their typical variation, as measured in study 1



thresholds are carefully determined and taking into account a “grey zone” of 3–6% in practice. These results further endorse the use of these thresholds to estimate the MMSS intensity. In contrast, our results did not justify the interchangeable use with local exercise thresholds ( $m[HHb]_{BP}$  and  $iEMG_{BP}$ ). Therefore, we do not recommend the stand-alone use of NIRS or EMG techniques when the purpose is to accurately determine the MMSS intensity from a single incremental exercise test.

**Acknowledgements** The authors would like to thank the subjects for their commitment to the study. This study was funded by the Special Research Fund of the Ghent University (Ghent, Belgium).

**Author contributions** KC, JB and JGB conceived and designed research. KC and ES conducted the experiments and analyzed the data. KC, JB and ES wrote the manuscript. All the authors revised and approved the manuscript.

**Funding** This research was funded by the Special Research Fund of the Ghent University (Ghent, Belgium) in the context of a doctoral fellowship (no. BOFDOC 2016003101).

**Availability of data and materials** The data that support the findings of this study are available from the corresponding author upon reasonable request.

**Code availability** Not applicable.

## Declarations

**Conflict of interest** Results of the present study are presented clearly, honestly and without fabrication, falsification or inappropriate data manipulation. No conflicts of interest, financial or otherwise, are declared by the authors.

**Ethics approval** This study was approved by the Ethical Committee of the Ghent University Hospital (Ghent, Belgium).

**Consent to participate** Informed consent for participation in this study was received from all subjects.

**Consent for publication** Consent for publication of the current study results was received from all subjects.

## References

- Beaver WL, Wasserman K, Whipp BJ (1986) A new method for detecting anaerobic threshold by gas exchange. *J Appl Physiol* 60:2020–2027. <https://doi.org/10.1152/jappl.1986.60.6.2020>
- Bellotti C, Calabria E, Capelli C, Pogliaghi S (2013) Determination of maximal lactate steady state in healthy adults: can NIRS help? *Med Sci Sports Exerc* 45:1208–1216. <https://doi.org/10.1249/MSS.0b013e3182828ab2>
- Bergstrom HC, Housh TJ, Zuniga JM et al (2014) Differences among estimates of critical power and anaerobic work capacity derived from five mathematical models and the three-minute all-out test. *J Strength Cond Res* 28:592–600. <https://doi.org/10.1519/JSC.0b013e31829b576d>
- Billat VL, Sirvent P, Py G et al (2003) The concept of maximal lactate steady state: a bridge between biochemistry, physiology and sport science. *Sports Med* 33:407–426. <https://doi.org/10.2165/00007256-200333060-00003>
- Binder RK, Wonisch M, Corra U et al (2008) Methodological approach to the first and second lactate threshold in incremental cardiopulmonary exercise testing. *Eur J Cardiovasc Prev Rehabil* 15:726–734. <https://doi.org/10.1097/HJR.0b013e328304fed4>
- Boone J, Vandekerckhove K, Coomans I et al (2016) An integrated view on the oxygenation responses to incremental exercise at the brain, the locomotor and respiratory muscles. *Eur J Appl Physiol* 116:2085–2102. <https://doi.org/10.1007/s00421-016-3468-x>
- Brickley G et al (2002) Physiological responses during exercise to exhaustion at critical power. *Eur J Appl Physiol* 88:146–151. <https://doi.org/10.1007/s00421-002-0706-1>
- Broxterman RM, Craig JC, Richardson RS (2018) The respiratory compensation point and the deoxygenation break point are not valid surrogates for critical power and maximum lactate steady state. *Med Sci Sports Exerc* 50:2379–2382. <https://doi.org/10.1249/MSS.0000000000001699>
- Bull AJ, Housh TJ, Johnson GO, Perry SR (2000) Effect of mathematical modeling on the estimation of critical power. *Med Sci Sports Exerc* 32:526. <https://doi.org/10.1097/00005768-20002000-00040>
- Burnley M, Jones AM (2007) Oxygen uptake kinetics as a determinant of sports performance. *Eur J Sport Sci* 7:63–79. <https://doi.org/10.1080/17461390701456148>
- Caen K (2021) Crossing the threshold : new insights on exercise thresholds and acute recovery from high-intensity exercise. Ghent University
- Caen K, Vermeire K, Bourgois JG, Boone J (2018) Exercise thresholds on trial: are they really equivalent? *Med Sci Sports Exerc* 50:1277–1284. <https://doi.org/10.1249/MSS.0000000000001547>
- Caen K, Boone J, Bourgois JG et al (2020) Translating ramp  $\dot{V}O_2$  into constant power output: a novel strategy that minds the gap. *Med Sci Sports Exerc*. <https://doi.org/10.1249/MSS.00000000000002328>
- Caen K, Pogliaghi S, Lievens M et al (2021) Ramp vs. step tests: valid alternatives to determine the maximal lactate steady-state intensity? *Eur J Appl Physiol*. <https://doi.org/10.1007/s00421-021-04620-9>
- Craig JC, Broxterman RM, Barstow TJ (2015) Considerations for identifying the boundaries of sustainable performance. *Med Sci Sports Exerc* 47:1997. <https://doi.org/10.1249/MSS.0000000000000676>
- Faude O, Kindermann W, Meyer T (2009) Lactate threshold concepts: how valid are they? *Sports Med* 39:469–490. <https://doi.org/10.2165/00007256-200939060-00003>
- Fontana FY, Keir DA, Bellotti C et al (2015) Determination of respiratory point compensation in healthy adults: can non-invasive near-infrared spectroscopy help? *J Sci Med Sport* 18:590–595. <https://doi.org/10.1016/j.jsams.2014.07.016>
- Galán-Rioja MÁ, González-Mohíno F, Poole DC, González-Ravé JM (2020) Relative proximity of critical power and metabolic/ventilatory thresholds: systematic review and meta-analysis. *Sports Med* 50:1771–1783. <https://doi.org/10.1007/s40279-020-01314-8>
- Heck H, Mader A, Hess G et al (1985) Justification of the 4-mmol/l lactate threshold. *Int J Sports Med* 6:117–130. <https://doi.org/10.1055/s-2008-1025824>
- Hermens HJ, Freriks B, Disselhorst-Klug C, Rau G (2000) Development of recommendations for SEMG sensors and sensor placement procedures. *J Electromyogr Kinesiol* 10:361–374. [https://doi.org/10.1016/S1050-6411\(00\)00027-4](https://doi.org/10.1016/S1050-6411(00)00027-4)

- Hopkins WG (2000) Measures of reliability in sports medicine and science. *Sports Med* 30:1–15. <https://doi.org/10.2165/00007256-200030010-00001>
- Iannetta D, Fontana FY, Maturana FM et al (2018) An equation to predict the maximal lactate steady state from ramp-incremental exercise test data in cycling. *J Sci Med Sport* 21:1274–1280. <https://doi.org/10.1016/j.jsams.2018.05.004>
- Iannetta D, Murias JM, Keir DA (2019) A simple method to quantify the V̇O<sub>2</sub> mean response time of ramp-incremental exercise. *Med Sci Sports Exerc* 51:1080–1086. <https://doi.org/10.1249/MSS.0000000000001880>
- Iannetta D, Inglis EC, Pogliaghi S et al (2020) A “step–ramp–step” protocol to identify the maximal metabolic steady state. *Med Sci Sports Exerc* 52:2011–2019. <https://doi.org/10.1249/MSS.0000000000002343>
- Inglis EC, Iannetta D, Keir DA, Murias JM (2020) Training-induced changes in the respiratory compensation point, deoxyhemoglobin break point, and maximal lactate steady state: evidence of equivalence. *Int J Sports Physiol Perform* 15:119–125. <https://doi.org/10.1123/ijsp.2019-0046>
- Jones AM, Burnley M, Black MI et al (2019) The maximal metabolic steady state: redefining the ‘gold standard.’ *Physiol Rep* 7:e14098. <https://doi.org/10.14814/phy2.14098>
- Keir DA, Murias JM, Paterson DH, Kowalchuk JM (2014) Breath-by-breath pulmonary O<sub>2</sub> uptake kinetics: effect of data processing on confidence in estimating model parameters: breath-by-breath pulmonary O<sub>2</sub> uptake kinetics. *Exp Physiol* 99:1511–1522. <https://doi.org/10.1113/expphysiol.2014.080812>
- Keir DA, Fontana FY, Robertson TC et al (2015a) Response. *Med Sci Sports Exerc* 47:1998–1999. <https://doi.org/10.1249/MSS.0000000000000677>
- Keir DA, Fontana FY, Robertson TC et al (2015b) Exercise intensity thresholds: identifying the boundaries of sustainable performance. *Med Sci Sports Exerc* 47:1932–1940. <https://doi.org/10.1249/MSS.0000000000000613>
- Keir DA, Paterson DH, Kowalchuk JM, Murias JM (2018a) Using ramp-incremental VO<sub>2</sub> responses for constant-intensity exercise selection. *Appl Physiol Nutr Metab* 43:882–892. <https://doi.org/10.1139/apnm-2017-0826>
- Keir DA, Pogliaghi S, Murias JM (2018b) The Respiratory compensation point and the deoxygenation break point are valid surrogates for critical power and maximum lactate steady state. *Med Sci Sports Exerc* 50:2375–2378. <https://doi.org/10.1249/MSS.0000000000001698>
- Keir DA, Pogliaghi S, Murias JM (2019) Response. *Med Sci Sports Exerc* 51:603. <https://doi.org/10.1249/MSS.0000000000001820>
- Koga S, Poole DC, Ferreira LF et al (2007) Spatial heterogeneity of quadriceps muscle deoxygenation kinetics during cycle exercise. *J Appl Physiol* 103:2049–2056. <https://doi.org/10.1152/jappphysiol.00627.2007>
- Koga S, Poole DC, Fukuoka Y et al (2011) Methodological validation of the dynamic heterogeneity of muscle deoxygenation within the quadriceps during cycle exercise. *Am J Physiol-Regul Integr Comp Physiol* 301:R534–R541. <https://doi.org/10.1152/ajpregu.00101.2011>
- Koo TK, Li MY (2016) A guideline of selecting and reporting intraclass correlation coefficients for reliability research. *J Chiropr Med* 15:155–163. <https://doi.org/10.1016/j.jcm.2016.02.012>
- Marwood S, Goulding RP, Roche DM (2019) Determining the upper limit of the metabolic steady state. *Med Sci Sports Exerc* 51:602. <https://doi.org/10.1249/MSS.0000000000001819>
- 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* 41:1197–1203. <https://doi.org/10.1139/apnm-2016-0248>
- Mattioni Maturana F, Fontana FY, Pogliaghi S et al (2018) Critical power: how different protocols and models affect its determination. *J Sci Med Sport* 21:742–747. <https://doi.org/10.1016/j.jsams.2017.11.015>
- Muniz-Pumares D, Karsten B, Triska C, Glaister M (2019) Methodological approaches and related challenges associated with the determination of critical power and curvature constant. *J Strength Cond Res* 33:584–596. <https://doi.org/10.1519/JSC.0000000000002977>
- Nagata A, Muro M, Moritani T, Yoshida T (1981) Anaerobic threshold determination by blood lactate and myoelectric signals. *Jpn J Physiol* 31:585–597. <https://doi.org/10.2170/jjphysiol.31.585>
- Poole DC, Ward SA, Gardner GW, Whipp BJ (1988) Metabolic and respiratory profile of the upper limit for prolonged exercise in man. *Ergonomics* 31:1265–1279. <https://doi.org/10.1080/00140138808966766>
- Poole DC, Burnley M, Vanhatalo A et al (2016) Critical power: an important fatigue threshold in exercise physiology. *Med Sci Sports Exerc* 48:2320–2334. <https://doi.org/10.1249/MSS.0000000000000939>
- Spencer MD, Murias JM, Paterson DH (2012) Characterizing the profile of muscle deoxygenation during ramp incremental exercise in young men. *Eur J Appl Physiol* 112:3349–3360. <https://doi.org/10.1007/s00421-012-2323-y>
- Takaishi T, Ono T, Yasuda Y (1992) Relationship between muscle fatigue and oxygen uptake during cycle ergometer exercise with different ramp slope increments. *Europ J Appl Physiol* 65:335–339. <https://doi.org/10.1007/BF00868137>
- Weir JP (2005) Quantifying test-retest reliability using the intraclass correlation coefficient and the SEM. *J Strength Cond Res* 19:231. <https://doi.org/10.1519/15184.1>
- Whipp BJ, Davis JA, Wasserman K (1989) Ventilatory control of the ‘isocapnic buffering’ region in rapidly-incremental exercise. *Respir Physiol* 76:357–367. [https://doi.org/10.1016/0034-5687\(89\)90076-5](https://doi.org/10.1016/0034-5687(89)90076-5)

**Publisher's Note** Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.