Progress Update and Challenges on $\dot{V}O_{2max}$ Testing and Interpretation

Marcos Martin-Rincon[1,2] and Jose A. L. Calbet[1,2,3*]

[1] Department of Physical Education, University of Las Palmas de Gran Canaria, Las Palmas de Gran Canaria, Spain, [2] Research Institute of Biomedical and Health Sciences (IUIBS), University of Las Palmas de Gran Canaria, Las Palmas de Gran Canaria, Spain, [3] Department of Physical Performance, The Norwegian School of Sport Sciences, Oslo, Norway

The maximal oxygen uptake ($\dot{V}O_{2max}$) is the primary determinant of endurance performance in heterogeneous populations and has predictive value for clinical outcomes and all-cause mortality. Accurate and precise measurement of $\dot{V}O_{2max}$ requires the adherence to quality control procedures, including combustion testing and the use of standardized incremental exercise protocols with a verification phase preceded by an adequate familiarization. The data averaging strategy employed to calculate the $\dot{V}O_{2max}$ from the breath-by-breath data can change the $\dot{V}O_{2max}$ value by 4–10%. The lower the number of breaths or smaller the number of seconds included in the averaging block, the higher the calculated $\dot{V}O_{2max}$ value with this effect being more prominent in untrained subjects. Smaller averaging strategies in number of breaths or seconds (less than 30 breaths or seconds) facilitate the identification of the plateau phenomenon without reducing the reliability of the measurements. When employing metabolic carts, averaging intervals including 15–20 breaths or seconds are preferable as a compromise between capturing the true $\dot{V}O_{2max}$ and identifying the plateau. In training studies, clinical interventions and meta-analysis, reporting of $\dot{V}O_{2max}$ in absolute values and inclusion of protocols and the averaging strategies arise as imperative to permit adequate comparisons. Newly developed correction equations can be used to normalize $\dot{V}O_{2max}$ to similar averaging strategies. A lack of improvement of $\dot{V}O_{2max}$ with training does not mean that the training program has elicited no adaptations, since peak cardiac output and mitochondrial oxidative capacity may be increased without changes in $\dot{V}O_{2max}$.

Keywords: calibration, cardiopulmonary exercise testing, ramp exercise, indirect calorimetry, reproducibility, breath-by-breath, metabolic cart, performance

INTRODUCTION

Relevance of $\dot{V}O_{2max}$ Determination in Performance and Health

The maximal oxygen uptake ($\dot{V}O_{2max}$) is the highest flow of oxygen (O_2) that can be used by the organism, representing the integrated capacity of the pulmonary, cardiovascular and muscle systems to take up, transport and utilize O_2. Classically, the determination of the $\dot{V}O_{2max}$ has been considered the gold standard for the assessment of the functional limits of the cardiorespiratory system (Taylor et al., 1955; Mitchell et al., 1958; Di Prampero, 2003; Levine, 2008). The $\dot{V}O_{2max}$ is one of the main factors associated with endurance performance (Joyner and Coyle, 2008) in various exercise modalities (Magel and Faulkner, 1967; Holmér et al., 1974; Maughan and Leiper, 1983). When a wide range of performances is included, $\dot{V}O_{2max}$ is the single best predictor of performance as shown in trained non-elite athletes, in whom $\dot{V}O_{2max}$ explained 90.2% of the total variance in a 16 km running time trial (McLaughlin et al., 2010). Although $\dot{V}O_{2max}$ does not predict performance in homogeneous groups of athletes (i.e., elite level), an exceptionally high $\dot{V}O_{2max}$ constitutes a prerequisite to compete at world-class level (Joyner and Coyle, 2008; Losnegard et al., 2013).

The terms cardiorespiratory fitness and $\dot{V}O_{2max}$ are used interchangeably, particularly in epidemiological studies. The assessment of $\dot{V}O_{2max}$ is gaining interest in clinical populations since it constitutes the strongest predictor of all-cause mortality when compared with other risk factors such as hypertension, smoking, obesity and diabetes (Myers et al., 2002; Lee et al., 2010). Furthermore, $\dot{V}O_{2max}$ is one of the objective criteria used to select candidates for heart transplantation (Mehra et al., 2016) and predicts the success of thoracic surgery (Brunelli et al., 2014). Thus, both in sport science and clinical medicine fields, there is a necessity for an accurate and precise determination of $\dot{V}O_{2max}$.

ASSESSMENT OF $\dot{V}O_{2max}$: EVOLUTION OF PROCEDURES AND THE PLATEAU PHENOMENON

Hill and Lupton (1923) are credited for performing the first assessments of $\dot{V}O_{2max}$. Years later, Taylor et al. (1955) proposed standardized procedures to determine $\dot{V}O_{2max}$, which were based on discontinuous protocols using three-min constant-intensity exercise periods carried out on subsequent days. At that time, $\dot{V}O_2$ measurements were based on the use of Douglas Bags and Tissot spirometers (Macfarlane, 2001) combined with Haldane's or Scholander's methods to obtain the concentration of O_2, which turned to be the gold-standard method for the assessment of $\dot{V}O_{2max}$. Taylor et al. (1955) defined the "plateau" in the $\dot{V}O_2$/intensity relationship as an increase in $\dot{V}O_2$ lower than 150 mL.min^{-1} or ~2 mL.kg^{-1}.min^{-1} with increasing exercise intensity. Continuous exercise protocols which allow for the assessment of additional variables of cardiorespiratory fitness were developed during the years to follow (Balke

and Ware, 1959; Bruce, 1971; Froelicher et al., 1974a,b). However, these continuous protocols often failed to fulfill Taylor's criteria for the plateauing effect (Froelicher et al., 1974a,b; Pollock et al., 1976).

The development of fast gas analyzers and flow sensors during the 1960s and 1970s together with the progressive miniaturization of computers resulted in the proposal of shorter continuous protocols which used the "limit of tolerance" of the subject (volitional exhaustion) as the criterion for test finalization (Whipp et al., 1981; Buchfuhrer et al., 1983). The duration of continuous protocols was judged as a critical variable for the achievement of "true" maximal $\dot{V}O_{2max}$ values (Buchfuhrer et al., 1983). These authors reported that durations between 8 and 17 min yielded the highest $\dot{V}O_{2max}$ values in five low to moderately trained subjects during treadmill and cycling ergometry, intuitively proposing 10 ± 2 min as the ideal duration. A critical factor that determines the length of the test is the rate at which intensity is increased over time or ramp slope, usually expressed in Watts/min (or speed/min). Research carried out in the last 40 years has confirmed that a similar $\dot{V}O_{2max}$ can be attained despite wide differences in the magnitude of the increments and duration of the tests (Zhang et al., 1991; Takaishi et al., 1992). Midgley et al. (2008) concluded that durations of 5 to 26 min could be optimal to achieve $\dot{V}O_{2max}$ in healthy individuals, provided that those with shorter durations are preceded by proper warm-up. Moreover, even repeated all-out sprint exercise has been shown to allow for the achievement of true $\dot{V}O_{2max}$ values, particularly when an active recovery is applied during the resting periods (Dorado et al., 2004; Gelabert-Rebato et al., 2018).

The utilization of gas exchange indirect calorimetry evolved rapidly from the Douglas Bag into technologies capable of examining the time course of gas exchange during exercise more comprehensively, namely mixing chambers and breath-by-breath systems. However, the Douglas Bag method is still considered as the gold-standard method (Shephard, 2017). Mixing-chamber automated systems measure expired gas fractions from several breaths collected into a small chamber (Wilmore and Costill, 1974; Wilmore et al., 1976), while breath-by-breath technology analyses the gas content of each breath by measuring instantaneous expiratory flow in conjunction with simultaneous tidal concentrations of exhaled gases (Beaver et al., 1973, 1981). Modern breath-by-breath metabolic carts are equipped with fast responding O_2 and CO_2 sensors and permit a precise and accurate assessment of the end-tidal O_2 and CO_2 pressures (Losa-Reyna et al., 2015). The latter can be used in conjunction with pulse oximetry for an indirect evaluation of pulmonary gas exchange, thus enhancing the utility of ergometry in clinical populations (Macfarlane, 2001; Sun et al., 2002).

Nevertheless, several technical limitations may undermine the precision and accuracy of breath-by-breath analyses if not adequately addressed (Di Prampero and Lafortuna, 1989). Yet, with state-of-the-art, well-calibrated and fast-responding metabolic carts, it is possible to closely match the levels of accuracy and precision of the Douglas bag method (Medbo et al., 2012; Nieman et al., 2013; Perez-Suarez et al., 2018; Martin-Rincon et al., 2019).

CONFIRMING THE ATTAINMENT OF $\dot{V}O_{2max}$: FROM THE PLATEAU PHENOMENON TO THE NECESSITY OF A VERIFICATION PHASE

Despite considerable technical improvements in metabolic carts, ergometers and exercise protocols, there is yet no universal agreement on the criteria for the attainment of a "true" $\dot{V}O_{2max}$. Until recent years, the most referenced criteria to confirm $\dot{V}O_{2max}$ attainment has been the presence of a plateau or leveling off in $\dot{V}O_2$, even though several investigations have shown that the incidence of the plateau is population-dependent (Achten et al., 2002; Huggett et al., 2005) and may not be manifested despite the attainment of a valid $\dot{V}O_{2max}$ (Day et al., 2003; Rossiter et al., 2006). In addition to the 150 mL or 2 mL.kg^{-1}.min^{-1} criterion previously proposed by Taylor et al. (1955) for the attainment of the plateau, other cut-off values have been postulated, as the $\Delta\dot{V}O_2 < 80$ mL.min^{-1} proposed by Astrand (1960) or the presence of a slope not different from zero with an increase in exercise intensity (Myers et al., 1990). However, the efficacy of these criteria to detect the plateau depends on the specific exercise protocol (Duncan et al., 1997), the population studied (Doherty et al., 2003; Beltrami et al., 2014), the subjects' experience (Gordon et al., 2015) and the breath-by-breath averaging strategy (Astorino, 2009; Nolan et al., 2014; Martin-Rincon et al., 2019). For example, attainment of a plateau may require to exercise with a marked recruitment of anaerobic metabolism during 1–3 min, entailing a high perception of effort and a strong central command to avoid task failure (Torres-Peralta et al., 2016a,b). Not surprisingly, subjects achieving a plateau present 4–5% higher levels of pulmonary ventilation (an indication of higher central command activation), respiratory exchange ratio (RER), and blood lactate concentration than those not reaching a plateau (Edvardsen et al., 2014). Given the limited duration of the plateau phase, its identification is facilitated by using shorter (15, 30 s) than longer (i.e., 60 s) time-averaging strategies (Astorino, 2009; Nolan et al., 2014).

This is further complicated by the influence that different exercise protocols and modes (Millet et al., 2009) may have on the highest $\dot{V}O_2$ value attainable (Midgley et al., 2008). Consequently, when uncertainty exists regarding whether a $\dot{V}O_{2max}$ value is real or not, the highest $\dot{V}O_2$ recorded is named $\dot{V}O_{2peak}$. Since many individuals reach a higher $\dot{V}O_{2max}$ value during a supramaximal rather than during an incremental exercise test, it is necessary to include a verification test few minutes after the end of the incremental test (Poole and Jones, 2017), even in clinical populations (Moreno-Cabanas et al., 2020).

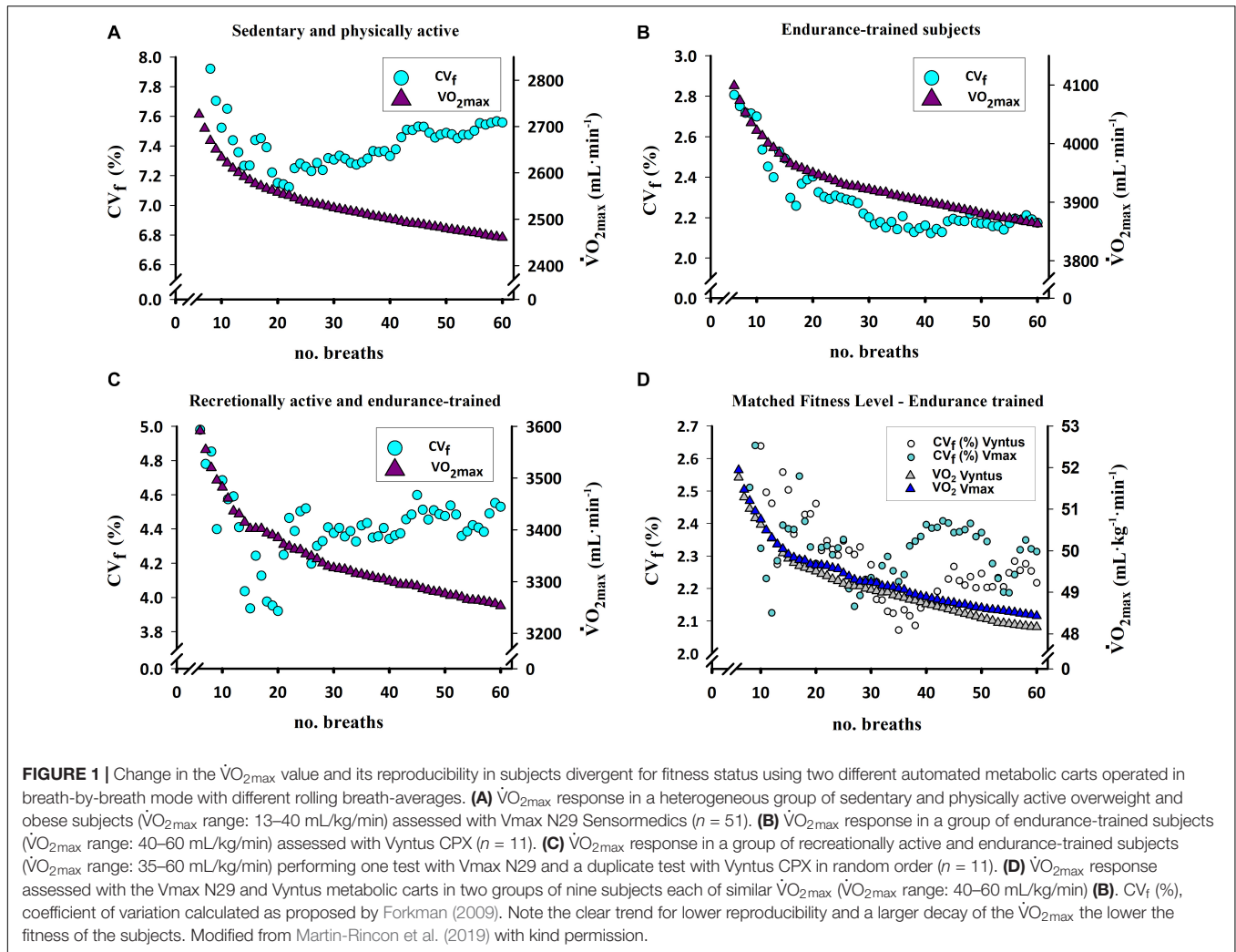## RECOMMENDED PROCEDURES FOR QUALITY CONTROL OF METABOLIC CARTS

Quality control procedures are mandatory to avoid technical errors. Most advanced metabolic carts are equipped with automated calibration routines that should be followed according to the recommendations of the manufacturers and using high-grade calibration O_2 and CO_2 gases. Besides the standard automated calibration, the flow sensors should be checked regularly with calibration syringes or commercially available metabolic simulators. Proper maintenance and replacement of flowmeters, O_2 cells, Nafion tubing and any other tubing and valves should follow the recommendations of the manufacturers. In addition, combustion tests of pure fuels such as methanol, butane or propane can be used to perform an integral check on the precision and accuracy of the metabolic carts. The combustion tests have the advantage of generating heat and some moisture, similar to the combustion present in living organisms (see, for example, Perez-Suarez et al., 2018). It is also advisable to perform regular checks at low and high exercise intensities, for example, on laboratory members holding a steady physical fitness level, as an additional biological check.

Verification of the performance of ergometers and treadmills is also necessary. In addition to the recommendations made by manufacturers, cross-validation biological checks with other ergometers are recommended. For example, repeated cardiorespiratory measurements at target intensities in the same person can be used to verify the stability of physiological variables. Treadmill speed and inclination can be easily verified manually or with odometer wheels.

IMPACT OF DATA PROCESSING STRATEGIES ON THE $\dot{V}O_2$ IMPUTED AS THE $\dot{V}O_{2max}$

Until recently, not much attention had been paid to the influence of the averaging interval or strategy on the $\dot{V}O_2$ value imputed as $\dot{V}O_{2max}$ and its reproducibility (Myers et al., 1990; Hill et al., 2003; Midgley et al., 2007; Astorino, 2009; Nolan et al., 2014; Smart et al., 2015; Dideriksen and Mikkelsen, 2017). Despite the inherent high variability between breaths, there is no universal consensus on how to average indirect calorimetry data (Robergs et al., 2010). We have recently demonstrated that the $\dot{V}O_2$ value imputed as $\dot{V}O_{2max}$ can fluctuate between 4 and 10% depending on the averaging strategy and the fitness levels (Martin-Rincon et al., 2019; **Figure 1**). We have shown that in subjects with a $\dot{V}O_{2max}$ lower than $\sim$40 mL/kg/min (13–40 mL/kg/min), the $\dot{V}O_{2max}$ value is $\sim$10% higher when using shorter (i.e., 10 breaths or seconds) averaging strategies rather than longer (i.e., 60 breaths or seconds in averaging block). This effect was slightly smaller ($\sim$6.5%) in higher fitness ($\dot{V}O_{2max}$ between 40 and 60 mL/kg/min) subjects. Similarly, from a short (i.e., 10 breaths or seconds) to an intermediate length in the averaging block (i.e., 30 breaths or seconds), the decrease is $\sim$7.5% in low-fit and $\sim$5% in those with higher fitness. Therefore, with shorter averaging strategies the $\dot{V}O_2$ value considered as $\dot{V}O_{2max}$ is larger, with this effect being more marked in untrained subjects, regardless of the use of breath- or time-based averaging strategies (Martin-Rincon et al., 2019). The time- and breath-averaged $\dot{V}O_2$ values produce similar results for a given length of the averaging block

FIGURE 1 | Change in the $\dot{V}O_{2max}$ value and its reproducibility in subjects divergent for fitness status using two different automated metabolic carts operated in breath-by-breath mode with different rolling breath-averages. **(A)** $\dot{V}O_{2max}$ response in a heterogeneous group of sedentary and physically active overweight and obese subjects ($\dot{V}O_{2max}$ range: 13–40 mL/kg/min) assessed with Vmax N29 Sensormedics ($n = 51$). **(B)** $\dot{V}O_{2max}$ response in a group of endurance-trained subjects ($\dot{V}O_{2max}$ range: 40–60 mL/kg/min) assessed with Vyntus CPX ($n = 11$). **(C)** $\dot{V}O_{2max}$ response in a group of recreationally active and endurance-trained subjects ($\dot{V}O_{2max}$ range: 35–60 mL/kg/min) performing one test with Vmax N29 and a duplicate test with Vyntus CPX in random order ($n = 11$). **(D)** $\dot{V}O_{2max}$ response assessed with the Vmax N29 and Vyntus metabolic carts in two groups of nine subjects each of similar $\dot{V}O_{2max}$ ($\dot{V}O_{2max}$ range: 40–60 mL/kg/min) **(B)**. CV$_f$ (%), coefficient of variation calculated as proposed by Forkman (2009). Note the clear trend for lower reproducibility and a larger decay of the $\dot{V}O_{2max}$ the lower the fitness of the subjects. Modified from Martin-Rincon et al. (2019) with kind permission.
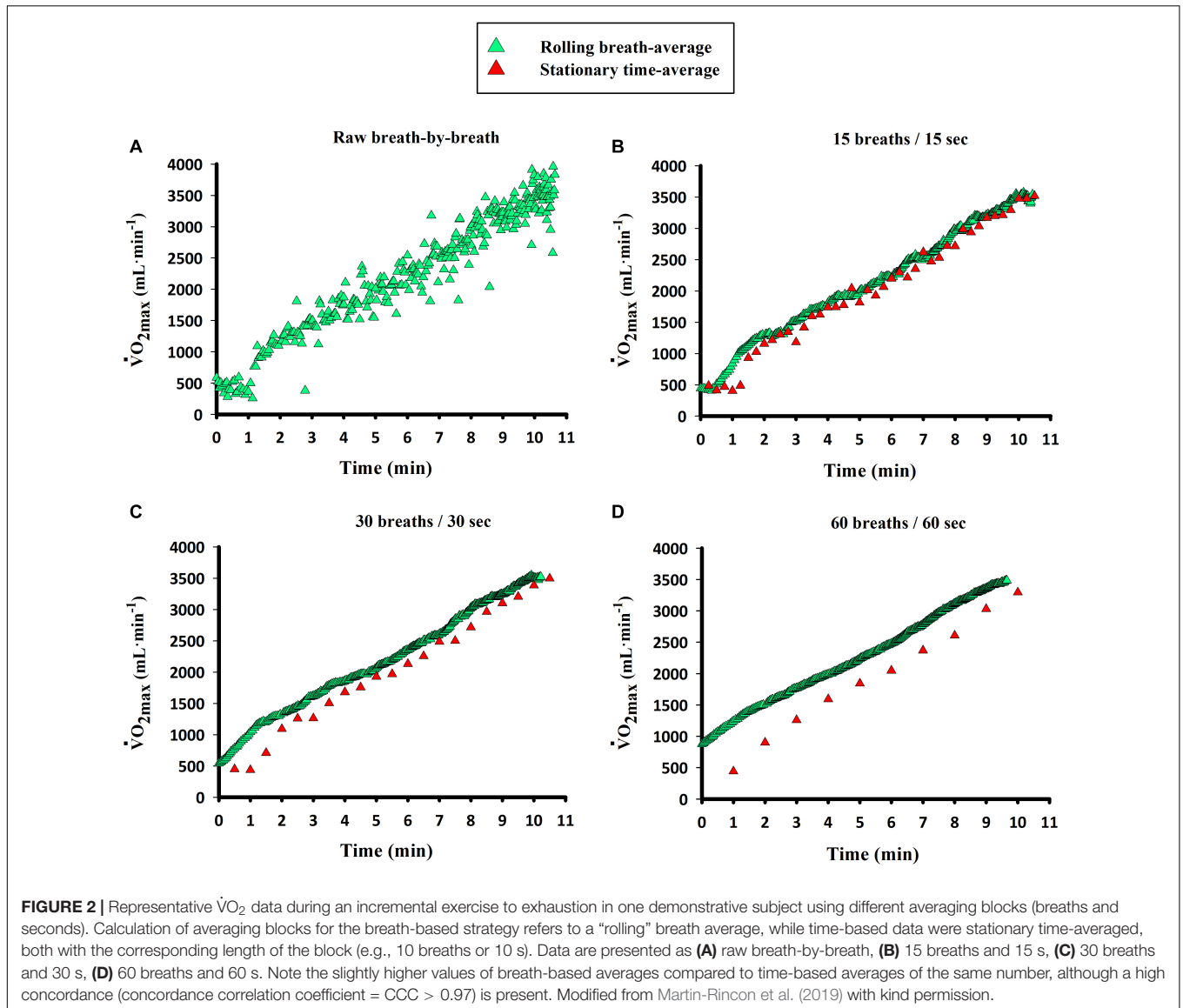
in breaths or seconds (Martin-Rincon et al., 2019; **Figure 2**). The values of $\dot{V}O_{2max}$ obtained with different averaging strategies can be interconverted with specific equations (Martin-Rincon et al., 2019). Given the fact that the $\dot{V}O_{2max}$ can be maintained only for a short time, the identification of the plateau is facilitated by short averaging intervals, as previously mentioned, without a negative impact of shorter averaging strategies on the reliability of the measurements (Martin-Rincon et al., 2019).

PHYSIOLOGICAL INTERPRETATION OF THE $\dot{V}O_{2max}$

The $\dot{V}O_{2max}$ value obtained in an exercise test should be checked against normative data (Edvardsen et al., 2013; Loe et al., 2013; Kaminsky et al., 2017), taking into consideration the exercise intensity achieved at the end of the test, exercise mode, age, sex, health status, fitness level, familiarization and the averaging strategy used to obtain the reference values. Since there is a close linear association between exercise intensity and $\dot{V}O_2$, there are several linear equations that can be used to predict the $\dot{V}O_2$

associated with a submaximal load as well as the $\dot{V}O_{2max}$ (Storer et al., 1990; Hall et al., 2004). The predicted or normative value can be confronted with the measured one, having in mind that departures up to 20% are possible (Malek et al., 2004). In the case of large differences between measured and estimated or reference values, assessment errors should be ruled out.

Since the $\dot{V}O_{2max}$ is a lumped parameter reflecting the flow of O_2 from the atmosphere to the mitochondria, the flow in downwards steps of the transport chain can never be higher than in the precedent step. $\dot{V}O_{2max}$ cannot be higher than O_2 delivery, and O_2 delivery cannot be higher than the O_2 flow in the lungs. For the correct physiological interpretation of the $\dot{V}O_{2max}$, it is critical to identify which the limiting step/factor is, holding in mind that the limiting factor may change with training (Zinner et al., 2016). In other words, it is essential to find out whether the limitation is mostly due to pulmonary ventilation, pulmonary gas exchange, cardiac output, muscle blood flow, arterial O_2 content, muscle O_2 diffusion capacity or the mitochondrial capacity to utilize O_2. In general, in healthy people, including elite athletes, the main limiting factor for $\dot{V}O_{2max}$ is O_2 delivery (Saltin and Calbet, 2006). Nevertheless, couch potatoes may

FIGURE 2 | Representative $\dot{V}O_2$ data during an incremental exercise to exhaustion in one demonstrative subject using different averaging blocks (breaths and seconds). Calculation of averaging blocks for the breath-based strategy refers to a "rolling" breath average, while time-based data were stationary time-averaged, both with the corresponding length of the block (e.g., 10 breaths or 10 s). Data are presented as **(A)** raw breath-by-breath, **(B)** 15 breaths and 15 s, **(C)** 30 breaths and 30 s, **(D)** 60 breaths and 60 s. Note the slightly higher values of breath-based averages compared to time-based averages of the same number, although a high concordance (concordance correlation coefficient = CCC > 0.97) is present. Modified from Martin-Rincon et al. (2019) with kind permission.

be limited by their capacity to use O_2, which can be readily tested by performing an incremental exercise test in hyperoxia (Cardus et al., 1998). Optimally, during the exercise test, the arterial hemoglobin saturation should be assessed indirectly by pulse oximetry. It is important to account for the impact that blood hemoglobin concentration may have on $\dot{V}O_{2max}$, and thus it is convenient to measure this variable before the tests (Calbet et al., 2002).

To facilitate the interpretation of results, the $\dot{V}O_{2max}$ should always be reported in absolute terms (i.e., $mL.min^{-1}$), together with the protocol used for its assessment and the averaging strategy. The $\dot{V}O_{2max}$ value adjusted for body weight has different meanings depending on the body composition, although the relative value (i.e., $\dot{V}O_{2max}$ in $mL.kg^{-1}.min^{-1}$) predicts better performance than the absolute value in competitions lasting from 60 to 90 s to several hours (Joyner and Coyle, 2008). In training studies and clinical interventions, the absolute value should

imperatively be reported such that a mechanistic explanation for the changes in $\dot{V}O_{2max}$, connected to the limiting factors, can be elaborated. Lack of improvement in $\dot{V}O_{2max}$ with training does not mean that the training program has failed or that it has elicited no adaptations. Actually, in elite athletes, peak cardiac output may be increased, as well as mitochondrial oxidative capacity, without changes in $\dot{V}O_{2max}$, due to an impairment of pulmonary gas exchange caused by the reduction of capillary blood mean transit time in the lung (Skattebo et al., 2020a). Moreover, peripheral adaptations may enhance the $\dot{V}O_{2peak}$ and O_2 extraction during small mass exercise without increasing the $\dot{V}O_{2max}$ measured during whole-body exercise (Skattebo et al., 2020b). Meta-analyses on the effects of an intervention on $\dot{V}O_{2max}$ should be based on absolute values, after taking into consideration the averaging strategies used in each study and the exercise modality, as previously discussed (Martin-Rincon et al., 2019).

FUTURE PERSPECTIVES

Although the $\dot{V}O_{2max}$ value varies depending on the averaging strategy, the fact that the values obtained with different averaging strategies are closely related allows the assumption that they hold a similar physiological meaning. However, the fact that $\dot{V}O_{2max}$ increases log-linearly with the shortening of the averaging interval (breath or time) (Martin-Rincon et al., 2019) implies that the real $\dot{V}O_{2max}$ value should be measured with a small number of breaths or a short time interval. Since the reproducibility is not significantly lowered by reducing the averaging interval between 6 and 60 breaths (or 6 and 60 s) (Martin-Rincon et al., 2019), the most appropriate approach would be using the shortest interval, since only one true $\dot{V}O_{2max}$ value must exist. From a physiological perspective it is not the same to compute the $\dot{V}O_{2max}$ value with a 6-s than a 60-s averaging strategy, since the latest may be 4–10% lower, depending on the fitness status. Given the relatively low trainability of $\dot{V}O_{2max}$ (Montero and Lundby, 2017), scientists and clinicians may easily misjudge the real condition of patients or the effects elicited by an intervention without accounting for the impact of the averaging strategy. Imputing an underestimated $\dot{V}O_{2max}$ has important implications for integrative physiology and pathophysiology. This is the case when comparing the $\dot{V}O_{2max}$ with maximal mitochondrial respiration values obtained *in vitro*, where the values of $\dot{V}O_{2max}$ used to represent the whole-body $\dot{V}O_{2max}$ will be smaller, and the excess mitochondrial respiratory capacity will be overestimated. Since the $\dot{V}O_{2max}$ is used to calculate cardiac output by the direct Fick method, for a given systemic arteriovenous O_2 (a-vO_2) difference, a $\dot{V}O_{2max}$ value obtained with a longer averaging interval will result in a lower calculated cardiac output by the Fick method. Likewise, the calculated pulmonary O_2 diffusing capacity would also be greater the higher the $\dot{V}O_{2max}$ imputed. Moreover, defining the $\dot{V}O_{2max}$ with a shorter averaging interval results in a value that can be maintained for a shorter time, affecting the procedures to determine the velocity or intensity at $\dot{V}O_{2max}$ (Billat and Koralsztein, 1996; Billat et al., 1999).

In the case that the $\dot{V}O_{2max}$ is used to calculate the systemic a-vO_2 difference using the indirect Fick method, attributing a higher $\dot{V}O_{2max}$ value would result in lower mixed venous O_2 contents and higher systemic O_2 extractions. Thus, when comparing the effect of different studies on variables determined by calculation from the $\dot{V}O_{2max}$, it is crucial to have into consideration the averaging strategy applied in each investigation. Therefore, these facts may contribute to explain some of the variability in the literature when secondary-outcome variables have been based on $\dot{V}O_{2max}$ values computed with large or small averaging strategies.

Based on the reproducibility of measurements, no particular sampling strategy seem superior for assessment of $\dot{V}O_{2max}$. Notwithstanding, given that the $\dot{V}O_{2max}$ can be sustained for a limited time, a shorter averaging strategy offers a higher probability for capturing the true $\dot{V}O_{2max}$, while facilitating the identification of the plateauing criteria (Martin-Rincon et al., 2019). Thus, when using modern metabolic carts, averaging intervals including those between 15 and 20 breaths (or seconds) are preferable as a compromise between capturing a $\dot{V}O_2$ value close to the true $\dot{V}O_{2max}$ and identifying the plateau. So far, the time resolution of the assessment of cardiac output and blood gases has not been sufficient as to clarify what factors determine the plateau phenomenon.

AUTHOR CONTRIBUTIONS

Both authors contributed similarly and approved the final version of the manuscript.

FUNDING

ACKNOWLEDGMENTS

REFERENCES

Achten, J., Gleeson, M., and Jeukendrup, A. E. (2002). Determination of the exercise intensity that elicits maximal fat oxidation. *Med. Sci. Sports Exerc.* 34, 92–97. doi: 10.1097/00005768-200201000-00015

Astorino, T. A. (2009). Alterations in VO$_{2max}$ and the VO$_2$ plateau with manipulation of sampling interval. *Clin. Physiol. Funct. Imaging* 29, 60–67. doi: 10.1111/j.1475-097X.2008.00835.x

Astrand, I. (1960). Aerobic work capacity in men and women with special reference to age. *Acta Physiol. Scand. Suppl.* 49, 1–92.

Balke, B., and Ware, R. W. (1959). An experimental study of physical fitness of air force personnel. *U.S. Armed Forces Med. J.* 10, 675–688.

Beaver, W. L., Lamarra, N., and Wasserman, K. (1981). Breath-by-breath measurement of true alveolar gas exchange. *J. Appl. Physiol.* 51, 1662–1675. doi: 10.1152/jappl.1981.51.6.1662

Beaver, W. L., Wasserman, K., and Whipp, B. J. (1973). On-line computer analysis and breath-by-breath graphical display of exercise function tests. *J. Appl. Physiol.* 34, 128–132. doi: 10.1152/jappl.1973.34.1.128

Beltrami, F. G., Wong Del, P., and Noakes, T. D. (2014). High prevalence of false-positive plateau phenomena during VO$_{2max}$ testing in adolescents. *J. Sci. Med. Sport* 17, 526–530. doi: 10.1016/j.jsams.2013.07.012

Billat, L. V., and Koralsztein, J. P. (1996). Significance of the velocity at VO$_{2max}$ and time to exhaustion at this velocity. *Sports Med.* 22, 90–108. doi: 10.2165/00007256-199622020-00004

Billat, V. L., Blondel, N., and Berthoin, S. (1999). Determination of the velocity associated with the longest time to exhaustion at maximal oxygen uptake. *Eur. J. Appl. Physiol.* 80, 159–161. doi: 10.1007/s004210050573

Bruce, R. A. (1971). Exercise testing of patients with coronary heart disease. Principles and normal standards for evaluation. *Ann. Clin. Res.* 3, 323–332.

Brunelli, A., Pompili, C., Salati, M., Refai, M., Berardi, R., Mazzanti, P., et al. (2014). Preoperative maximum oxygen consumption is associated with prognosis after pulmonary resection in stage I non-small cell lung cancer. *Ann. Thorac. Surg.* 98, 238–242. doi: 10.1016/j.athoracsur.2014.04.029

Buchfuhrer, M. J., Hansen, J. E., Robinson, T. E., Sue, D. Y., Wasserman, K., and Whipp, B. J. (1983). Optimizing the exercise protocol for cardiopulmonary

assessment. *J. Appl. Physiol.* 55, 1558–1564. doi: 10.1152/jappl.1983.55.5.1558

Calbet, J. A., Radegran, G., Boushel, R., Sondergaard, H., Saltin, B., and Wagner, P. D. (2002). Effect of blood haemoglobin concentration on VO_{2max} and cardiovascular function in lowlanders acclimatised to 5260 m. *J. Physiol.* 545, 715–728. doi: 10.1113/jphysiol.2002.029108

Cardus, J., Marrades, R. M., Roca, J., Barbera, J. A., Diaz, O., Masclans, J. R., et al. (1998). Effects of $F_{I}O_2$ on leg VO_2 during cycle ergometry in sedentary subjects. *Med. Sci. Sports Exerc.* 30, 697–703. doi: 10.1097/00005768-199805000-00009

Day, J. R., Rossiter, H. B., Coats, E. M., Skasick, A., and Whipp, B. J. (2003). The maximally attainable VO_2 during exercise in humans: the peak vs. maximum issue. *J. Appl. Physiol.* 95, 1901–1907. doi: 10.1152/japplphysiol.00024.2003

Di Prampero, P. E. (2003). Factors limiting maximal performance in humans. *Eur. J. Appl. Physiol.* 90, 420–429. doi: 10.1007/s00421-003-0926-z

Di Prampero, P. E., and Lafortuna, C. L. (1989). Breath-by-breath estimate of alveolar gas transfer variability in man at rest and during exercise. *J. Physiol.* 415, 459–475. doi: 10.1113/jphysiol.1989.sp017731

Dideriksen, K., and Mikkelsen, U. R. (2017). Reproducibility of incremental maximal cycle ergometer tests in healthy recreationally active subjects. *Clin. Physiol. Funct. Imaging* 37, 173–182. doi: 10.1111/cpf.12283

Doherty, M., Nobbs, L., and Noakes, T. D. (2003). Low frequency of the "plateau phenomenon" during maximal exercise in elite British athletes. *Eur. J. Appl. Physiol.* 89, 619–623. doi: 10.1007/s00421-003-0845-z

Dorado, C., Sanchis-Moysi, J., and Calbet, J. A. (2004). Effects of recovery mode on performance, O_2 uptake, and O_2 deficit during high-intensity intermittent exercise. *Can. J. Appl. Physiol.* 29, 227–244. doi: 10.1139/h04-016

Duncan, G. E., Howley, E. T., and Johnson, B. N. (1997). Applicability of VO_{2max} criteria: discontinuous versus continuous protocols. *Med. Sci. Sports Exerc.* 29, 273–278. doi: 10.1097/00005768-199702000-00017

Edvardsen, E., Hansen, B. H., Holme, I. M., Dyrstad, S. M., and Anderssen, S. A. (2013). Reference values for cardiorespiratory response and fitness on the treadmill in a 20- to 85-year-old population. *Chest* 144, 241–248. doi: 10.1378/chest.12-1458

Edvardsen, E., Hem, E., and Anderssen, S. A. (2014). End criteria for reaching maximal oxygen uptake must be strict and adjusted to sex and age: a cross-sectional study. *PLoS One* 9:e85276. doi: 10.1371/journal.pone.0085276

Forkman, J. (2009). Estimator and tests for common coefficients of variation in normal distributions. *Commun. Stat. Theory Methods* 38, 233–251. doi: 10.1080/03610920802187448

Froelicher, V. F., Brammell, H., Davis, G., Noguera, I., Stewart, A., and Lancaster, M. C. (1974a). A comparison of the reproducibility and physiologic response to three maximal treadmill exercise protocols. *Chest* 65, 512–517. doi: 10.1378/chest.65.5.512

Froelicher, V. F., Brammell, H., Davis, G., Noguera, I., Stewart, A., and Lancaster, M. C. (1974b). A comparison of three maximal treadmill exercise protocols. *J. Appl. Physiol.* 36, 720–725. doi: 10.1152/jappl.1974.36.6.720

Gelabert-Rebato, M., Wiebe, J. C., Martin-Rincon, M., Gericke, N., Perez-Valera, M., Curtelin, D., et al. (2018). *Mangifera indica* L. Leaf extract in combination with luteolin or quercetin enhances VO_2 peak and peak power output, and preserves skeletal muscle function during ischemia-reperfusion in humans. *Front. Physiol.* 9:740. doi: 10.3389/fphys.2018.00740

Gordon, D., Caddy, O., Merzbach, V., Gernigon, M., Baker, J., Scruton, A., et al. (2015). Prior knowledge of trial number influences the incidence of plateau at VO_{2max}. *J. Sports Sci. Med.* 14, 47–53.

Hall, C., Figueroa, A., Fernhall, B., and Kanaley, J. A. (2004). Energy expenditure of walking and running: comparison with prediction equations. *Med. Sci. Sports Exerc.* 36, 2128–2134. doi: 10.1249/01.mss.0000147584.87788.0e

Hill, A. V., and Lupton, H. (1923). Muscular exercise, lactic acid, and the supply and utilization of oxygen. *Q. J. Med.* 16, 135–171. doi: 10.1093/qjmed/os-16.62.135

Hill, D. W., Stephens, L. P., Blumoff-Ross, S. A., Poole, D. C., and Smith, J. C. (2003). Effect of sampling strategy on measures of VO_2peak obtained using commercial breath-by-breath systems. *Eur. J. Appl. Physiol.* 89, 564–569. doi: 10.1007/s00421-003-0843-1

Holmér, I., Lundin, A., and Eriksson, B. O. (1974). Maximum oxygen uptake during swimming and running by elite swimmers. *J. Appl. Physiol.* 36, 711–714. doi: 10.1152/jappl.1974.36.6.711

Huggett, D. L., Connelly, D. M., and Overend, T. J. (2005). Maximal aerobic capacity testing of older adults: a critical review. *J. Gerontol. A Biol. Sci. Med. Sci.* 60, 57–66. doi: 10.1093/gerona/60.1.57

Joyner, M. J., and Coyle, E. F. (2008). Endurance exercise performance: the physiology of champions. *J. Physiol.* 586, 35–44. doi: 10.1113/jphysiol.2007.143834

Kaminsky, L. A., Imboden, M. T., Arena, R., and Myers, J. (2017). Reference standards for cardiorespiratory fitness measured with cardiopulmonary exercise testing using cycle ergometry: data from the fitness registry and the importance of exercise national database (friend) registry. *Mayo Clin. Proc.* 92, 228–233. doi: 10.1016/j.mayocp.2016.10.003

Lee, D.-C., Artero, E. G., Sui, X., and Blair, S. N. (2010). Mortality trends in the general population: the importance of cardiorespiratory fitness. *Int. J. Neuropsychopharmacol.* 24, 27–35. doi: 10.1177/1359786810382057

Levine, B. D. (2008). VO_{2max}: what do we know, and what do we still need to know? *J. Physiol.* 586, 25–34. doi: 10.1113/jphysiol.2007.147629

Loe, H., Rognmo, O., Saltin, B., and Wisloff, U. (2013). Aerobic capacity reference data in 3816 healthy men and women 20-90 years. *PLoS One* 8:e64319. doi: 10.1371/journal.pone.0064319

Losa-Reyna, J., Torres-Peralta, R., Henriquez, J. J., and Calbet, J. A. (2015). Arterial to end-tidal PCO_2 difference during exercise in normoxia and severe acute hypoxia: importance of blood temperature correction. *Physiol. Rep.* 3:e12512. doi: 10.14814/phy2.12512

Losnegard, T., Myklebust, H., Spencer, M., and Hallen, J. (2013). Seasonal variations in VO_{2max}, O_2-cost, O_2-deficit, and performance in elite cross-country skiers. *J. Strength Cond. Res.* 27, 1780–1790. doi: 10.1519/JSC.0b013e31827368f6

Macfarlane, D. J. (2001). Automated metabolic gas analysis systems: a review. *Sports Med.* 31, 841–861. doi: 10.2165/00007256-200131120-00002

Magel, J. R., and Faulkner, J. A. (1967). Maximum oxygen uptakes of college swimmers. *J. Appl. Physiol.* 22, 929–933. doi: 10.1152/jappl.1967.22.5.929

Malek, M. H., Berger, D. E., Housh, T. J., Coburn, J. W., and Beck, T. W. (2004). Validity of VO_{2max} equations for aerobically trained males and females. *Med. Sci. Sports Exerc.* 36, 1427–1432. doi: 10.1249/01.mss.0000135795.60449.ce

Martin-Rincon, M., Gonzalez-Henriquez, J. J., Losa-Reyna, J., Perez-Suarez, I., Ponce-Gonzalez, J. G., De La Calle-Herrero, J., et al. (2019). Impact of data averaging strategies on VO_{2max} assessment: mathematical modeling and reliability. *Scand. J. Med. Sci. Sports* 29, 1473–1488. doi: 10.1111/sms.13495

Maughan, R. J., and Leiper, J. B. (1983). Aerobic capacity and fractional utilisation of aerobic capacity in elite and non-elite male and female marathon runners. *Eur. J. Appl. Physiol.* 52, 80–87. doi: 10.1007/BF00429030

McLaughlin, J. E., Howley, E. T., Bassett, D. R. Jr., Thompson, D. L., and Fitzhugh, E. C. (2010). Test of the classic model for predicting endurance running performance. *Med. Sci. Sports Exerc.* 42, 991–997. doi: 10.1249/MSS.0b013e3181c0669d

Medbo, J. I., Mamen, A., and Beltrami, F. G. (2012). Examination of the moxus modular metabolic system by the douglas-bag technique. *Appl. Physiol. Nutr. Metab.* 37, 860–871. doi: 10.1139/h2012-056

Mehra, M. R., Canter, C. E., Hannan, M. M., Semigran, M. J., Uber, P. A., Baran, D. A., et al. (2016). The 2016 international society for heart lung transplantation listing criteria for heart transplantation: a 10-year update. *J. Heart Lung Transplant.* 35, 1–23. doi: 10.1016/j.healun.2015.10.023

Midgley, A. W., Bentley, D. J., Luttikholt, H., Mcnaughton, L. R., and Millet, G. P. (2008). Challenging a dogma of exercise physiology: does an incremental exercise test for valid VO_{2max} determination really need to last between 8 and 12 minutes? *Sports Med.* 38, 441–447. doi: 10.2165/00007256-200838060-00001

Midgley, A. W., Mcnaughton, L. R., and Carroll, S. (2007). Effect of the VO_2 time-averaging interval on the reproducibility of O_{2max} in healthy athletic subjects. *Clin. Physiol. Funct. Imaging* 27, 122–125. doi: 10.1111/j.1475-097x.2007.00725.x

Millet, G. P., Vleck, V. E., and Bentley, D. J. (2009). Physiological differences between cycling and running: lessons from triathletes. *Sports Med.* 39, 179–206. doi: 10.2165/00007256-200939030-00002

Mitchell, J. H., Sproule, B. J., and Chapman, C. B. (1958). The physiological meaning of the maximal oxygen intake test. *J. Clin. Invest.* 37, 538–547. doi: 10.1172/jci103636

Montero, D., and Lundby, C. (2017). Refuting the myth of non-response to exercise training: 'non-responders' do respond to higher dose of training. *J. Physiol.* 595, 3377–3387. doi: 10.1113/jp273480

Moreno-Cabanas, A., Ortega, J. F., Morales-Palomo, F., Ramirez-Jimenez, M., and Mora-Rodriguez, R. (2020). Importance of a verification test to accurately assess VO_{2max} in unfit individuals with obesity. *Scand. J. Med. Sci. Sports* 30, 583–590. doi: 10.1111/sms.13602

Myers, J., Prakash, M., Froelicher, V., Do, D., Partington, S., and Atwood, J. E. (2002). Exercise capacity and mortality among men referred for exercise testing. *N. Engl. J. Med.* 346, 793–801. doi: 10.1056/NEJMoa011858

Myers, J., Walsh, D., Sullivan, M., and Froelicher, V. (1990). Effect of sampling on variability and plateau in oxygen uptake. *J. Appl. Physiol.* 68, 404–410. doi: 10.1152/jappl.1990.68.1.404

Nieman, D. C., Austin, M. D., Dew, D., and Utter, A. C. (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, 136–145. doi: 10.1080/15438627.2012.757227

Nolan, P. B., Maddison, R., and Dalleck, L. (2014). The incidence of VO_2 plateau at VO_{2max} in a cardiac-diseased population. *Int. J. Sports Med.* 35, 118–124. doi: 10.1055/s-0033-1345135

Perez-Suarez, I., Martin-Rincon, M., Gonzalez-Henriquez, J. J., Fezzardi, C., Perez-Regalado, S., Galvan-Alvarez, V., et al. (2018). Accuracy and precision of the COSMED K5 portable analyser. *Front. Physiol.* 9:1764. doi: 10.3389/fphys.2018.01764

Pollock, M. L., Bohannon, R. L., Cooper, K. H., Ayres, J. J., Ward, A., White, S. R., et al. (1976). A comparative analysis of four protocols for maximal treadmill stress testing. *Am. Heart J.* 92, 39–46. doi: 10.1016/s0002-8703(76)80401-2

Poole, D. C., and Jones, A. M. (2017). Measurement of the maximum oxygen uptake VO_{2max}: VO_2peak is no longer acceptable. *J. Appl. Physiol.* 122, 997–1002. doi: 10.1152/japplphysiol.01063.2016

Robergs, R. A., Dwyer, D., and Astorino, T. (2010). Recommendations for improved data processing from expired gas analysis indirect calorimetry. *Sports Med.* 40, 95–111. doi: 10.2165/11319670-000000000-00000

Rossiter, H. B., Kowalchuk, J. M., and Whipp, B. J. (2006). A test to establish maximum O_2 uptake despite no plateau in the O_2 uptake response to ramp incremental exercise. *J. Appl. Physiol.* 100, 764–770. doi: 10.1152/japplphysiol.00932.2005

Saltin, B., and Calbet, J. A. L. (2006). Counterpoint: in health and in a normoxic environment, VO_{2max} is not limited primarily by cardiac output and locomotor muscle blood flow - Rebuttal. *J. Appl. Physiol.* 100, 747–748.

Shephard, R. J. (2017). Open-circuit respirometry: a brief historical review of the use of Douglas bags and chemical analyzers. *Eur. J. Appl. Physiol.* 117, 381–387. doi: 10.1007/s00421-017-3556-6

Skattebo, O., Calbet, J. A. L., Rud, B., Capelli, C., and Hallen, J. (2020a). Contribution of oxygen extraction fraction to maximal oxygen uptake in healthy young men. *Acta Physiol.* 4:e13486. doi: 10.1111/apha.13486

Skattebo, O., Capelli, C., Rud, B., Auensen, M., Calbet, J. A. L., and Hallen, J. (2020b). Increased oxygen extraction and mitochondrial protein expression after small muscle mass endurance training. *Scand. J. Med. Sci. Sports*. doi: 10.1111/sms.13707 [Epub ahead of print].

Smart, N. A., Jeffriess, L., Giallauria, F., Vigorito, C., Vitelli, A., Maresca, L., et al. (2015). Effect of duration of data averaging interval on reported peak VO_2 in patients with heart failure. *Int. J. Cardiol.* 182, 530–533. doi: 10.1016/j.ijcard.2014.12.174

Storer, T. W., Davis, J. A., and Caiozzo, V. J. (1990). Accurate prediction of VO_{2max} in cycle ergometry. *Med. Sci. Sports Exerc.* 22, 704–712. doi: 10.1249/00005768-199010000-00024

Sun, X. G., Hansen, J. E., Oudiz, R. J., and Wasserman, K. (2002). Gas exchange detection of exercise-induced right-to-left shunt in patients with primary pulmonary hypertension. *Circulation* 105, 54–60. doi: 10.1161/hc0102.101509

Takaishi, T., Ono, T., and Yasuda, Y. (1992). Relationship between muscle fatigue and oxygen uptake during cycle ergometer exercise with different ramp slope increments. *Eur. J. Appl. Physiol.* 65, 335–339. doi: 10.1007/BF00868137

Taylor, H. L., Buskirk, E., and Henschel, A. (1955). Maximal oxygen intake as an objective measure of cardio-respiratory performance. *J. Appl. Physiol.* 8, 73–80. doi: 10.1152/jappl.1955.8.1.73

Torres-Peralta, R., Losa-Reyna, J., Morales-Alamo, D., Gonzalez-Izal, M., Perez-Suarez, I., Ponce-Gonzalez, J. G., et al. (2016a). Increased P_IO_2 at exhaustion in hypoxia enhances muscle activation and swiftly relieves fatigue: A Placebo or a P_IO_2 dependent effect? *Front. Physiol.* 7:333. doi: 10.3389/fphys.2016.00333

Torres-Peralta, R., Morales-Alamo, D., Gonzalez-Izal, M., Losa-Reyna, J., Perez-Suarez, I., Izquierdo, M., et al. (2016b). Task failure during exercise to exhaustion in normoxia and hypoxia is due to reduced muscle activation caused by central mechanisms while muscle metaboreflex does not limit performance. *Front. Physiol.* 6:414. doi: 10.3389/fphys.2015.00414

Whipp, B. J., Davis, J. A., Torres, F., and Wasserman, K. (1981). A test to determine parameters of aerobic function during exercise. *J. Appl. Physiol.* 50, 217–221. doi: 10.1152/jappl.1981.50.1.217

Wilmore, J. H., and Costill, D. L. (1974). Semiautomated systems approach to the assessment of oxygen uptake during exercise. *J. Appl. Physiol.* 36, 618–620. doi: 10.1152/jappl.1974.36.5.618

Wilmore, J. H., Davis, J. A., and Norton, A. C. (1976). An automated system for assessing metabolic and respiratory function during exercise. *J. Appl. Physiol.* 40, 619–624. doi: 10.1152/jappl.1976.40.4.619

Zhang, Y. Y., Johnson, M. C. II, Chow, N., and Wasserman, K. (1991). Effect of exercise testing protocol on parameters of aerobic function. *Med. Sci. Sports Exerc.* 23, 625–630.

Zinner, C., Morales-Alamo, D., Ortenblad, N., Larsen, F. J., Schiffer, T. A., Willis, S. J., et al. (2016). The physiological mechanisms of performance enhancement with sprint interval training differ between the upper and lower extremities in humans. *Front. Physiol.* 7:426. doi: 10.3389/fphys.2016.00426