The Oxygen Uptake Plateau—A Critical Review of the Frequently Misunderstood Phenomenon

Max Niemeyer1 · Raphael Knaier2,3 · Ralph Beneke1

Accepted: 8 April 2021 / Published online: 29 April 2021
© The Author(s) 2021

Abstract
A flattening of the oxygen uptake–work rate relationship at severe exercise indicates the achievement of maximum oxygen uptake (\( \dot{V}O_2 \text{max} \)). Unfortunately, a distinct plateau (\( \dot{V}O_2 \text{pl} \)) at \( \dot{V}O_2 \text{max} \) is not found in all participants. The aim of this investigation was to critically review the influence of research methods and physiological factors on the \( \dot{V}O_2 \text{pl} \) incidence. It is shown that many studies used inappropriate definitions or methodical approaches to check for the occurrence of a \( \dot{V}O_2 \text{pl} \). In contrast to the widespread assumptions it is unclear whether there is higher \( \dot{V}O_2 \text{pl} \) incidence in (uphill) running compared to cycling exercise or in discontinuous compared to continuous incremental exercise tests. Furthermore, most studies that evaluated the validity of supramaximal verification phases, reported verification bout durations, which are too short to ensure that \( \dot{V}O_2 \text{max} \) have been achieved by all participants. As a result, there is little evidence for a higher \( \dot{V}O_2 \text{pl} \) incidence and a corresponding advantage for the diagnoses of \( \dot{V}O_2 \text{max} \) when incremental tests are supplemented by supramaximal verification bouts. Preliminary evidence suggests that the occurrence of a \( \dot{V}O_2 \text{pl} \) in continuous incremental tests is determined by physiological factors like anaerobic capacity, \( \dot{V}O_2 \)-kinetics and accumulation of metabolites in the submaximal intensity domain. Subsequent studies should take more attention to the use of valid \( \dot{V}O_2 \text{pl} \) definitions, which require a cut-off at ~50% of the submaximal \( \dot{V}O_2 \) increase and rather large sampling intervals. Furthermore, if verification bouts are used to verify the achievement of \( \dot{V}O_2 \text{peak} / \dot{V}O_2 \text{max} \), it should be ensured that they can be sustained for sufficient durations.

Key Points

A near-constant \( \dot{V}O_2 \) despite a further increase in work rate (= \( \dot{V}O_2 \text{pl} \)) indicates attainment of \( \dot{V}O_2 \text{max} \); however, not all participants demonstrate a \( \dot{V}O_2 \text{pl} \) at the end of an incremental test.

There is lack of convincing evidence to show that the incidence of the \( \dot{V}O_2 \text{pl} \) is influenced by exercise mode, exercise protocol, aerobic fitness, anthropometrics, or age.

Preliminary evidence indicates that a fast \( \dot{V}O_2 \)-kinetics as well as a high anaerobic capacity and anaerobic threshold related to \( \dot{V}O_2 \text{max} \) seem to increase the chance that a plateau at \( \dot{V}O_2 \text{max} \) occurs.
1 Introduction

The first description of the oxygen uptake plateau (\(\dot{V}O_2\text{pl}\)) is attributed to Archibald Vivian Hill [1–3]. The former Nobel Prize laureate and his colleague Hartley Lupton collected their breathing gases with a Douglas bag while running on flat terrain with varying running speeds. In one participant they found that \(\dot{V}O_2\) increased with running speed, but remained constant for speeds beyond 260 m per minute. Hill and Lupton [3] interpreted the flattening of the \(\dot{V}O_2\)-running-speed-relationship as an indication that maximum oxygen uptake (\(\dot{V}O_2\text{max}\)) has been achieved. In subsequent studies the flattening of the \(\dot{V}O_2\)-running-speed-relationship or -work-rate-relationship in the severe-intensity domain, as shown in Fig. 1, has been called “\(\dot{V}O_2\)-plateau” or “levelling off” [4–7]. Nowadays, \(\dot{V}O_2\text{max}\) is considered as one of the most important measurements in exercise physiology and sports medicine [8–10]. However, it has been often ignored that one can be certain that \(\dot{V}O_2\text{max}\) has been reached only if \(\dot{V}O_2\) remains more or less constant despite an increase in work rate. Thus, the diagnosis of \(\dot{V}O_2\text{max}\) requires—per definition—a \(\dot{V}O_2\text{pl}\), regardless whether the plateau occurs at the end of a continuous incremental test, between the final stages of a discontinuous test or between an incremental test and a subsequent verification test (see Sect. 2.3), as highlighted by Poole and Jones [11].

Unfortunately, only a fraction of participants that undergo exercise testing shows a \(\dot{V}O_2\text{pl}\). The reported incidences vary between 17 and 94%, even in larger studies with more than 50 participants, especially depending on the used \(\dot{V}O_2\text{pl}\) definition (see Sect. 2.1) [7, 12–19]. This led to extensive discussions concerning the concept and diagnosis of \(\dot{V}O_2\text{max}\) throughout the last three decades [2, 11, 20–30]. The aim of the present investigation is to critically review the existing body of literature about the \(\dot{V}O_2\text{pl}\) and aspects of \(\dot{V}O_2\text{max}\) diagnosis with special respect to methodological and physiological determinants of the incidence of the \(\dot{V}O_2\text{pl}\).

2 Methodological Determinants

Although described for the first time almost a century ago, there is still a lot of confusion about the methodological determinants and the correct diagnosis of the \(\dot{V}O_2\text{pl}\) phenomenon [23, 24, 27]. This led to the assumption that the \(\dot{V}O_2\text{pl}\) is a calculation artefact rather than an indication of a physiological event [20, 24, 31–33]. In contrast, other authors suggested that the absence of a \(\dot{V}O_2\text{pl}\) is mostly caused by inappropriate exercise modes and protocols or data analysing approaches [11, 14, 34–36]. Potential effects of the data analysing approach, exercise mode and protocol on the \(\dot{V}O_2\text{pl}\) incidence will be critically revisited below.
2.1 Plateau Definitions

The first quantitative approach for the diagnosis of the \( \dot{V}O_{2\text{pl}} \) was provided by Taylor et al. [7]. They performed discontinuous exercise tests on a treadmill with 3-min durations at a speed of 3.13 m s\(^{-1}\). Treadmill inclination was increased by 2.5\% from test to test at subsequent testing days. When the increase in \( \dot{V}O_2 \) between subsequent testing days was less than 150 mL min\(^{-1}\) (\( -2.1 \) mL min\(^{-1}\) kg\(^{-1}\)) a \( \dot{V}O_{2\text{pl}} \) was assumed. This cut-off had been determined based on corresponding treadmill tests in the submaximal intensity domain i.e. two or more steps before a \( \dot{V}O_{2\text{pl}} \) occurred in a subgroup of 13 participants. They found an average increase in \( \dot{V}O_2 \) of 299.3 ± 86.5 mL min\(^{-1}\) (\( -4.2 \pm 1.1 \) mL min\(^{-1}\) kg\(^{-1}\)) with a range of 159–470 mL min\(^{-1}\) (\( -2.2–5.9 \) mL min\(^{-1}\) kg\(^{-1}\)) per 2.5\% grade steps. Based on this finding the authors [7] concluded that if the \( \dot{V}O_2 \) increases for less than 150 mL min\(^{-1}\) or 2.1 mL min\(^{-1}\) kg\(^{-1}\): “…there is small chance of making an error in deciding that the maximal oxygen intake had been reached”.

The cut-offs of Taylor et al. [7] are still widely used for \( \dot{V}O_{2\text{pl}} \) diagnoses [23, 25], despite this assumption has been challenged in subsequent studies (see Sect. 2.3.2). But more importantly, these cut-offs had been validated for specific testing conditions only. They were deduced from a discontinuous exercise protocol with bouts of constant running speeds long enough to reflect the complete fast component of steady-state \( \dot{V}O_2 \)-kinetics and a specific increase in oxygen demand per subsequent running bout [37, 38]. Instead of such highly specific testing conditions others combined the Taylor et al. [7] cut-offs with arbitrary time intervals, also called sampling intervals, during various one-time continuous incremental test irrespective of the test-specific \( \dot{V}O_2 \)-response [14, 32, 34, 39–41]. Such modifications may lead to mismatches between presumptions implied by Taylor et al. [7] and test-specific increases in \( \dot{V}O_2 \) per time or workload increment induced using other testing protocols [38, 42]. For example, sampling intervals of 30 s were widely used to check whether a \( \dot{V}O_{2\text{pl}} \) occurs or not [14, 40, 43–45]. This means that the average \( \dot{V}O_2 \) of the last and next-to-last 30 s period of a test were compared. At a ramp test with an incremental rate of 30 W min\(^{-1}\) the mean difference in work rate between adjacent 30 s sampling intervals is 15 W leading to a mean increase in \( \dot{V}O_2 \) between adjacent sampling intervals in the submaximal intensity domain of \( \approx 150 \) mL min\(^{-1}\) [42, 46]. Here application of the 150 mL min\(^{-1}\) cut-off by Taylor et al. [7] may lead to \( \dot{V}O_{2\text{pl}} \) diagnoses despite no flattening at \( \dot{V}O_{2\text{max}} \) occurs. Consequently, application of such fixed or absolute cut-offs on arbitrarily selected testing protocols and time intervals may provoke high risks of false \( \dot{V}O_{2\text{pl}} \) diagnoses, as described by Beltrami et al. [32] and Marsh [38]. To avoid such mismatches, relative cut-offs which consider the increase in \( \dot{V}O_2 \) in the submaximal intensity domain of a specific testing protocol are strongly recommended [37, 38]. However, even when a relative cut-off had been used several \( \dot{V}O_{2\text{pl}} \)’s in the submaximal intensity domain had been found by Beltrami et al. [32]. Based on this finding they concluded that the \( \dot{V}O_{2\text{pl}} \) may reflect a calculation artefact rather than an indication of a true physiological event.

Mathematically a plateau is defined as part of a function with a slope equal to zero. However, the breath-by-breath \( \dot{V}O_2 \) response shows large variability, which is mainly caused by irregularities in the rate and depth of ventilation [25, 47]. Consequently, even under steady-state conditions, the \( \dot{V}O_2 \) values of subsequent sampling intervals are never exactly similar (i.e. slope \( \neq 0 \)) as shown by Myers et al. [47]. Therefore, if the cut-off is set at a slope equal to zero, potential \( \dot{V}O_{2\text{pl}} \)’s which are caused by the limitation of the body to transport and utilise \( O_2 \), may remain undetected (false-negative \( \dot{V}O_{2\text{pl}} \) diagnosis). To avoid such false-negative \( \dot{V}O_{2\text{pl}} \) diagnoses the cut-off must be set at a slope or value which is higher than the ventilation-induced variability of \( \dot{V}O_2 \) [42].

On the other hand, the variability of \( \dot{V}O_2 \) may result in plateau-like \( \dot{V}O_2 \)-responses in the submaximal intensity domain of incremental exercise tests, as shown in Fig. 2. Such plateau-like \( \dot{V}O_2 \)-responses may result in an erroneously levelling-off-diagnosis if the test shown in Fig. 2 had been terminated prematurely, for example, due to a lack of motivation or pain tolerance, at a work rate of \( \approx 285 \) W. Therefore, plateaus in the submaximal intensity domain can be classified as false-positive \( \dot{V}O_{2\text{pl}} \) diagnoses [32, 42]. To be able to discriminate between plateaus that are simply caused by the variability of ventilation (false-positive \( \dot{V}O_{2\text{pl}} \)’s) and those which are caused by the limitation of the body to transport or utilise \( O_2 \) (real \( \dot{V}O_{2\text{pl}} \)’s), a plateau at \( \dot{V}O_{2\text{max}} \) must be more pronounced than potential plateaus occurring in the submaximal intensity domain [42].

To reduce the risk of false-positive \( \dot{V}O_{2\text{pl}} \) diagnoses several studies used rather restrictive cut-offs which were set at considerably less than 50\% of the mean increase or difference in \( \dot{V}O_2 \) in the submaximal intensity domain [6, 43–45]. Since the variability of \( \dot{V}O_2 \) shows a Gaussian distribution [42, 47, 48] and such restrictive cut-offs may be additionally less than the accuracy of common breath-by-breath devices this approach increases the risk of false-negative \( \dot{V}O_{2\text{pl}} \) diagnoses [2, 42, 47]. Based on a recent study [42] this becomes clear. The authors determined the risk of false-positive and false-negative \( \dot{V}O_{2\text{pl}} \) diagnoses for different work rate defined sampling intervals at a ramp test with an incremental rate of 30 W min\(^{-1}\). They started with a sampling interval of 30 W (= 60 s), which means that several slopes were fitted
into different 30 W intervals in the submaximal intensity domain to check for the rate/risk of false-positive diagnoses. At this sampling interval the risk of false-positive diagnoses was 12.7%. Since the variability of the slopes was Gaussian distributed it can be calculated that the risk of false-negative diagnoses is also 12.7% if the cut-off is set at half of the increase in \(\dot{V}O_2\) in the submaximal intensity domain [42]. If the cut-off is set instead of at one-third of the increase in \(\dot{V}O_2\) in the submaximal intensity domain, the risk of false-positive \(\dot{V}O_2\)-plateaus reduces to 6.4%. However, this goes along with a more pronounced increase in the risk of false-negative \(\dot{V}O_2\)-plateaus to 22.4%, such that the combined risk of false-positive and false-negative \(\dot{V}O_2\)-plateaus is slightly higher than if the cut-off set at half of the increase (28.8% vs. 25.4%). This demonstrates that the cut-off should be set at approximately half of the increase in the submaximal intensity domain to enable an equal risk of false-positive and false-negative \(\dot{V}O_2\)-plateaus diagnoses [42].

However, a combined risk of false \(\dot{V}O_2\)-plateaus diagnoses of ~25% is quite high and do not allow to detect a real \(\dot{V}O_2\)-plateau with sufficient certainty. Therefore, the sampling interval was subsequently increased to 40 (= 80 s) and 50 W (= 100 s), which led to a reduction of the combined risk of false-positive and false-negative \(\dot{V}O_2\)-plateaus diagnoses to ~12.7% [42]. Application of the latter \(\dot{V}O_2\)-plateau definition to the final 50 W increment of the ramp tests resulted in the detection of a considerably higher \(\dot{V}O_2\)-plateau incidence (35.7%) than the expected rate of false-positive \(\dot{V}O_2\)-plateaus due to the variability of \(\dot{V}O_2\) (1.6%) [42]. This indicates that a \(\dot{V}O_2\)-plateau in terms of a real change in the \(\dot{V}O_2\)/P-ratio at \(\dot{V}O_2\)-max exists which is not simply caused by the variability of ventilation or a calculation artefact as suggested by some researchers [20, 24, 31–33, 47]. However, the observed \(\dot{V}O_2\)-plateau incidence at maximal intensity of ~35% in the Niemeyer et al. [42] study was much lower than the corresponding incidences of 90–100% reported occasionally [14, 34]. Since the latter studies used very short sampling intervals combined with inappropriate absolute cut-offs these extremely high \(\dot{V}O_2\)-plateau incidences likely reflect a high fraction of false-positive \(\dot{V}O_2\)-plateaus diagnoses [32, 38, 42]. Therefore, assumptions [14, 34] that the absence of a \(\dot{V}O_2\)-plateau is caused by an inappropriate plateau definition and/or insufficient data analyses methodology appear unfounded. \(\dot{V}O_2\)-plateau incidences of about 20–60% seem realistic if more appropriate \(\dot{V}O_2\)-plateau definitions are used [13, 18, 19, 42, 49–52].

With respect to the written above, an appropriate \(\dot{V}O_2\)-plateau definition requires (1) that the cut-off is set at approximately 50% of the increase in \(\dot{V}O_2\) in the submaximal intensity domain of a specific testing protocol; (2) that rather large sampling intervals are used to check for the occurrence of a \(\dot{V}O_2\)-plateau. For ramp exercise tests with incremental rates of 30 ± 10 W \text{ min}^{-1} sampling intervals of approximately 40–50 W seems to be ideal [42]. This means that the increase in \(\dot{V}O_2\) during the final 40–50 W of an incremental exercise test must be less than half of the increase in the submaximal intensity domain to be certain that a real \(\dot{V}O_2\)-plateau occurs [42]. Importantly, this approach cannot be transferred to incremental tests with much lower incremental rates which are usually applied to sedentary or clinical populations, as highlighted by Niemeyer et al. [42]. The same applies to other exercise modes, such as rowing or running.
As a result, it is impossible to evaluate the validity of the plateau definitions of all published studies which are relevant for the present review with sufficient certainty. This applies also to studies that did not report sampling intervals or incremental rates. Therefore, we opted against to exclude studies that used \( \text{VO}_{2\text{pl}} \) definitions with unclear validity. Instead, potential effects of the used \( \text{VO}_{2\text{pl}} \) definitions on the corresponding findings in the subsequent sections are discussed. Studies that used rather large sampling intervals (i.e. the \( \text{VO}_{2\text{pl}} \) was determined from more than the final 30 W or 60 s) and adequate cut-offs were classified as “probably valid”. An adequate cut-off was assumed when the cut-off was derived from the increase in \( \text{VO}_2 \) in the submaximal intensity domain or set at approximately half of the expected increase in \( \text{VO}_2 \) in the submaximal intensity domain.

### 2.2 Exercise Mode

It is well-known that \( \text{VO}_2 \text{max} / \text{VO}_{2\text{peak}} \) varies between exercise modes. Untrained or non-specifically trained participants achieve a 5–15% higher \( \text{VO}_2 \text{max} / \text{VO}_{2\text{peak}} \) in uphill treadmill running than in cycling or other exercise modes [5, 53–56]. Only specifically trained athletes may reach higher or comparable \( \text{VO}_2 \text{max} / \text{VO}_{2\text{peak}} \) in their accustomed discipline than in (uphill) running [56, 57]. Based on the assumption that a \( \text{VO}_{2\text{pl}} \) occurs only at the highest rate of oxygen uptake that participants can attain when the mode fits perfectly to their abilities (= \( \text{VO}_2 \text{max} \)), it is widely believed that there is a higher \( \text{VO}_{2\text{pl}} \) incidence in (uphill) running or in the specifically trained discipline compared to other exercise modes [27, 35, 36, 58]. However, only two studies analysed \( \text{VO}_{2\text{pl}} \) incidences at different exercise modes based on within-subject designs [39, 44]. Both studies described significantly higher \( \text{VO}_{2\text{pl}} \) incidences in uphill treadmill running (~50%) than in cycling (8% and 20%, respectively). However, both studies used absolute cut-offs for the diagnoses of a \( \text{VO}_{2\text{pl}} \) and did not account for different increases in \( \text{VO}_2 \) between exercise modes and specific testing protocols, respectively. For example, Gordon et al. [44] used an incremental rate of 0.5 W s\(^{-1}\) for cycling leading to an increase in \( \text{VO}_2 \) in the submaximal intensity domain between consecutive 30 s sampling intervals of about ~150 mL min\(^{-1}\) [42]. In contrast, the incremental rate of 0.5% per 30 s-interval at a running speed of 10 km h\(^{-1}\) for the treadmill test led to an increase in \( \text{VO}_2 \) of about ~60 mL min\(^{-1}\) between consecutive 30 s sampling intervals. Since the cut-off was set at 50 mL min\(^{-1}\) for both exercise modes, there was a much higher risk of false-positive \( \text{VO}_{2\text{pl}} \) diagnoses at the treadmill compared to the cycling tests. A similar problem applies to the study by Rivera-Brown et al. [39]. Consequently, based on the present evidence it is unclear whether there is really a higher \( \text{VO}_{2\text{pl}} \) incidence in (uphill) running compared to other exercise modes.

Irrespective of this, it has been shown that participants may demonstrate a clear \( \text{VO}_{2\text{pl}} \) at the end of incremental cycling tests despite a higher \( \text{VO}_2 \text{max} \) during treadmill or combined leg and arm cycling tests [59–61]. Therefore, a \( \text{VO}_{2\text{pl}} \) indicates only the achievement of a mode- or task-specific \( \text{VO}_2 \text{max} \) and not the achievement of the highest rate of oxygen uptake that a participant can attain when most of the skeletal muscles are activated and/or the exercise mode corresponds to the training specificity [61].

### 2.3 Exercise Protocol

#### 2.3.1 Exercise Duration and Incremental Rate

The effect of different incremental rates and the resulting time to exhaustion on \( \text{VO}_2 \text{max} / \text{VO}_{2\text{peak}} \) has been observed in several studies. Some of them suggest that \( \text{VO}_2 \text{max} / \text{VO}_{2\text{peak}} \) can be measured in a wide range of incremental rates leading to exhaustion between approximately 5 and 25 min [62, 63]. However, other studies showed that an incremental rate should be used, which leads to exhaustion between 8 and 12 or 16 min [64–67]. To the best of our knowledge, there is only one study with a within-subject design which checked the effect of incremental rate on the \( \text{VO}_{2\text{pl}} \) incidences [64]. In this study incremental rates between 15 and 90 W min\(^{-1}\) were applied to 16 trained participants (eight men and eight women) to accomplish ramp test durations of approximately 5, 8, 12 and 16 min. The authors [64] did not find any systematic effects of the test durations or related incremental rates to the incidence of the \( \text{VO}_{2\text{pl}} \). The highest incidences of the \( \text{VO}_{2\text{pl}} \) were found at ramp tests of 8 and 16 min duration. Unfortunately, the authors [64] defined a \( \text{VO}_{2\text{pl}} \) as an increase in \( \text{VO}_2 \) of <50 mL min\(^{-1}\) within the final 30 s disregarding the fact that the expected increases in \( \text{VO}_2 \) in the submaximal intensity domain ranged from 75 mL min\(^{-1}\) to 450 mL min\(^{-1}\) between the different protocols. Thus, a much higher risk of false-positive \( \text{VO}_{2\text{pl}} \) diagnoses in those protocols with lower incremental rates has to be expected (see Sect. 2.1).

The combined findings from studies using either high or low incremental rates indicate that \( \text{VO}_{2\text{pl}} \) occurs in a wide range of incremental rates and corresponding test durations. Thus, \( \text{VO}_{2\text{pl}} \) were reported in incremental tests leading to exhaustion within ~8 but also ~24 min [49, 68–70]. This indicates that \( \text{VO}_{2\text{pl}} \) occur within the same range of incremental rates and resulting test durations which enables the achievement of \( \text{VO}_2 \text{max} \). However, further studies using appropriate plateau definitions are needed to verify whether there is an optimal incremental rate or test duration for the \( \text{VO}_{2\text{pl}} \) occurrence.
2.3.2 Continuous vs. Discontinuous Exercise

Early studies used discontinuous exercise tests to measure \( \dot{\text{VO}}_2 \) \(_{\text{max}} \) \([3, 6, 7]\). As shown in Fig. 3 discontinuous tests are characterised by several discrete constant load exercises, which are separated by long resting periods (i.e. hours or days). For prior applications of discontinuous protocols used to measure \( \dot{\text{VO}}_2 \) \(_{\text{max}} \), the intensity of the constant load bouts was increased until the corresponding increase in \( \dot{\text{VO}}_2 \) fell short of a predefined cut-off or until the work rate could not be sustained for a predefined duration \([5–7, 52]\). Due to their time-consuming nature and improvements in measurement equipment, discontinuous protocols have been replaced by continuous incremental exercise protocols.

As shown in several studies \( \dot{\text{VO}}_2 \) \(_{\text{max}} / \dot{\text{VO}}_2 \) \(_{\text{peak}} \) does not differ between discontinuous and continuous exercise tests \([52, 53, 55, 71]\). However, compared to the discontinuous test protocol used by Taylor et al. \([7]\) \( \dot{\text{VO}}_2 \) \(_{\text{pl}} \) incidences are considerably lower in most studies using continuous tests \([12, 13, 15, 18, 19, 51]\). This led to the assumption that \( \dot{\text{VO}}_2 \) \(_{\text{pl}} \) incidences are lower in continuous compared to discontinuous exercise tests \([2, 11, 13]\). To the best of our knowledge, there is only one study that compared \( \dot{\text{VO}}_2 \) \(_{\text{pl}} \) incidences of continuous and discontinuous exercise tests using a within-subject design \([52]\). In this study the original discontinuous exercise protocol of Taylor et al. \([7]\), and a continuous incremental treadmill test with an increase of treadmill grade of 2.5% every minute, were used to measure \( \dot{\text{VO}}_2 \) \(_{\text{max}} \). In contrast to the study of Taylor et al. \([7]\), Duncan et al. \([52]\) increased the treadmill grade not only until the increase in \( \dot{\text{VO}}_2 \) between subsequent treadmill grades was < 2.1 mL min\(^{-1}\) kg\(^{-1}\), but until the participants could not sustain the discontinuous exercises for 3 min. Analysing the final two 2.5% grade steps with the cut-off defined by Taylor et al. \([7]\), Duncan et al. \([52]\) found a \( \dot{\text{VO}}_2 \) \(_{\text{pl}} \) incidence of 60% and 50% at the discontinuous and continuous tests, respectively. However, if they accepted a \( \dot{\text{VO}}_2 \) \(_{\text{pl}} \) as soon as the increase in \( \dot{\text{VO}}_2 \) between consecutive 2.5% discontinuous steps was less than the cut-off as implemented by Taylor et al. \([7]\) the \( \dot{\text{VO}}_2 \) \(_{\text{pl}} \) incidence in the discontinuous test was 80%. This finding is supported by an early study \([5]\) which showed that the original approach of Taylor et al. \([7]\) leads to several false-positive \( \dot{\text{VO}}_2 \) \(_{\text{pl}} \) diagnoses. Also, other studies \([58, 72, 73]\) using discontinuous exercise tests identified far lower incidences of \( \dot{\text{VO}}_2 \) \(_{\text{pl}} \) than Taylor et al. \([7]\), however, those numbers were in the range of continuous test studies \([18, 19, 42, 50, 51]\). In conclusion, there is no convincing evidence for a higher \( \dot{\text{VO}}_2 \) \(_{\text{pl}} \) incidence in discontinuous compared to continuous exercise tests. The extremely high \( \dot{\text{VO}}_2 \) \(_{\text{pl}} \) incidences described by Taylor et al. \([7]\) are likely caused by the specific approach of that study.
which was likely to result in a high rate of false-positive \( \dot{V}O_{2pl} \) diagnoses [5, 52]. Since the only study [52] that compared the \( \dot{V}O_{2pl} \) incidence between discontinuous and continuous tests with a within-subject design had a rather small sample size (\( n = 10 \)) further studies are needed to verify whether the \( \dot{V}O_{2pl} \) incidence differs between discontinuous and continuous exercise protocols.

### 2.3.3 Combination of Continuous and Discontinuous Exercise: The Verification Phase

Based on low \( \dot{V}O_{2pl} \) incidences and questionable validity of secondary exhaustion criteria the use of \( \dot{V}O_{2max} \) verification tests/phases are strongly recommended by some researchers [11, 74]. As shown in Fig. 4 this approach includes a constant-load exercise test, which is performed after a common continuous incremental load test and a recovery phase. If \( \dot{V}O_{2peak} \) of the incremental and verification phases do not differ despite the verification bout was performed on a higher work rate than maximum work rate of the incremental test i.e. supramaximal work rate (related to maximum work rate at termination of the incremental test), a \( \dot{V}O_{2pl} \) occurs and \( \dot{V}O_{2max} \) of the incremental test is considered as verified [11].

If \( \dot{V}O_{2peak} \) of the verification phase is higher than \( \dot{V}O_{2peak} \) of the incremental test no indication of a \( \dot{V}O_{2pl} \) has been verified and \( \dot{V}O_{2max} \) of the incremental test is considered as falsified [11]. Several studies reported considerable lower plateau occurrences during the final part of a single incremental test as compared to the difference between the \( \dot{V}O_{2peak} \) values of an incremental test and the \( \dot{V}O_{2peak} \) values of a subsequent supramaximal verification bout [69, 75–79].

However, besides a supramaximal work rate of the verification bout [11], there is another important prerequisite for valid \( \dot{V}O_{2pl} \) diagnoses based on the combination of incremental and verification phases. Because of a rather slow \( \dot{V}O_{2} \)-kinetics, the duration of the verification bout must be long enough to reach \( \dot{V}O_{2max} \) [11, 28]. According to Hill et al. [80], the minimum duration of an exhaustive exercise trial to rise \( \dot{V}O_{2}\) to its maximum value is 2.3 ± 0.3 min. A more detailed study of Caputo and Denadai [81] showed that in untrained and non-specifically trained participants durations of at least 3.5 ± 0.5 and 2.8 ± 0.5 min are required, respectively. In contrast, in specifically endurance-trained athletes durations of 2.0 ± 0.5 min were sufficiently long to rise \( \dot{V}O_{2}\) to its maximum value [81]. These minimum durations

![Fig. 4 Oxygen uptake (\( \dot{V}O_{2} \)) (above) and work rate (below) profile of an incremental ramp test, which was followed by a recovery and a verification phase. Note that peak oxygen uptake (\( \dot{V}O_{2peak} \)) of the ramp and verification phases were nearly similar despite the verification bout was performed at a 5% higher work rate than maximum work rate of the ramp bout, which indicates the achievement of a \( \dot{V}O_{2pl} \). \( \Delta \dot{V}O_{2peak} \): difference between the peak oxygen uptake values achieved at the incremental and verification phase. \( \dot{V}O_{2pl} \) was calculated as the mean of the highest 30 s-interval of the incremental and verification bout, respectively.](https://example.com/fig4.png)
are supported by VO$_2$ kinetic studies, which showed that endurance-trained participants have considerable shorter time constants of VO$_2$ kinetics compared to non-specific or untrained individuals [82–84]. Thus, to construct a real VO$_2$pl between incremental and verification phases, the work rate of the verification bout must be at supramaximal load and needs to be sustained for a minimum of ~2 (trained), ~3 (non-specifically trained) or ~3.5 (untrained) minutes. Since heavy/severe prior exercise leads to speeding of VO$_2$ kinetics [85, 86] it seems to be likely that VO$_2$max will be achieved on average slightly earlier in verification bouts, which are performed a few minutes (up to ~30 min) after incremental tests. However, these numbers represent the mean values of the corresponding cohorts, which mean that some participants need even longer to achieve VO$_2$max. These durations should be therefore viewed as an absolute minimum, even if the verification bout is performed in a primed state.

Table 1 summarises published incremental rates, VO$_2$max, and time-to-exhaustions (TTE) of studies which evaluated the use of supramaximal verification tests and reported exact TTE values [37, 68–70, 76, 77, 79, 87–94]. Based on this table it becomes clear that only three out of the 22 studies or sub-studies reported mean TTE values of verification bouts of sufficient duration to achieve VO$_2$max with convincing probability [70, 77, 87]. The study reporting the longest TTE performed a supramaximal verification bout at 105% of the work rate at the termination of a common incremental test and a recovery phase of at least 24 h in a group of untrained participants [87]. The TTE of 4.2 ± 2.0 was more than 100% higher than the mean TTE of all other cycling studies performing a verification bout at 105% (see Table 1). However, the large standard deviation of the TTE of the verification bout indicates that a considerable fraction of participants did not reach the required minimum test duration for untrained participants of >3:30 min.

The two other studies that reported mean TTE values of sufficient duration were performed on treadmills [70, 77]. One of them [70] used an incremental test with a rather low incremental rate which led to a more than twice as large incremental test durations (23.9 ± 2.1 min) compared to most other studies (see Table 1). In incremental tests the maximally achievable work rate or running speed is positively related to the incremental rate [62, 67]. Therefore, the work rate of a verification bout which is performed at a given percentage of maximum work rate or running speed depends on the incremental rate of the preceding incremental test. At a first glance, this seems to support the use of low incremental rates to guarantee that subsequent verification bouts can be sustained for sufficient durations. However, this benefit comes with the risk to reduce the chance that VO$_2$max will be achieved in the incremental phase as the optimal duration for VO$_2$max testing may be exceeded [64–67]. Interestingly, the study of Sanchez-Otero et al. [70], described a significantly lower VO$_2$peak in the verification bout compared to the incremental test despite the verification bout lasted long enough. This is potentially caused by the phenomenon that well-trained athletes do not reach VO$_2$max by fast component VO$_2$ response during square-wave exercise [95–97], which further limits the validity of supramaximal verification testing to increase the VO$_2$pl incidence.

Nevertheless, most other studies reported similar mean VO$_2$peak-values in the incremental and verification phases (see Table 1). However, this does not indicate that a VO$_2$pl and VO$_2$max were reached in every single participant as demonstrated by the following example. Midgley et al. [77] did not find a significant difference in mean VO$_2$peak values between the incremental and verification tests (Table 1). Analysing the 32 incremental tests, only 16 VO$_2$pl were found. In contrast, VO$_2$max of 26 out of the 32 incremental tests was verified by a not considerably higher VO$_2$ value (<2%) in the verification bout, suggesting a VO$_2$pl incidence between the incremental and the verification bout of 81%. However, they [77] did not consider that in seven out of the 26 positive verification bouts the VO$_2$ was more than 3% lower than the corresponding VO$_2$peak value of the incremental test. Midgley et al. [77] attributed this lower verification bout VO$_2$ values to insufficient verification bout durations. Therefore, it is unclear whether VO$_2$max was achieved in the correspondence incremental tests. When subtracting these seven tests only 19 out of the 32 incremental tests (59%) were really verified by the verification bouts, which is close to the number of identified VO$_2$pl’s via incremental tests only (50%). The latter example combined with the fact that most studies in Table 1 reported even shorter TTE values of the verification bouts indicate rather limited evidence for an advantage of supramaximal verification tests over single incremental test procedures in the determination of a VO$_2$pl and VO$_2$max.

In contrast, it has been shown that submaximal verification bouts (e.g. 95% of maximum work rate) may be useful to identify submaximal VO$_2$peak values of incremental tests, especially in sedentary [98] and clinical [99] populations, but also in recreationally active participants [42]. However, it is important to note that a submaximal verification bout does not allow constructing a VO$_2$pl, since this is defined as a more or less constant VO$_2$ despite an increase in work rate [7, 11]. As a result, one cannot be certain that VO$_2$max has been achieved when the VO$_2$peak of a submaximal verification bout is similar to VO$_2$peak of an incremental test. However, if the VO$_2$peak of a submaximal verification bout is higher than the VO$_2$peak of an incremental test one can be certain that VO$_2$max has not been reached during the incremental phase. Furthermore, the higher VO$_2$peak of the submaximal verification bout is closer to VO$_2$max. Therefore, submaximal verification bouts, which can be sustained for
### Table 1  Summary of the most important findings of studies or sub-studies that evaluated the validity of supramaximal $\dot{V}O_2$ max verification testing in healthy participants sorted by time to exhaustion (TTE) of the verification bout

| Authors (year) | Subjects (n) | Exercise mode | Incremental test | Verification test |
|----------------|--------------|---------------|------------------|-------------------|
|                |              |               | Incremental rate | Intensity (%$P_{max}$) | TTE (min) | $\dot{V}O_2$ max (L min$^{-1}$/ mL min$^{-1}$ kg$^{-1}$) |
|                 |              |               | TTE (min) | $\dot{V}O_2$ max (L min$^{-1}$/ mL min$^{-1}$ kg$^{-1}$) | Recovery duration |               |
| Astorino et al. [87] | Untrained (n = 15) | Cycling | 14–21 (W min$^{-1}$) | 14–21 (W min$^{-1}$) | 9.2 ± 2.4 | 2.37 ± 0.69 | ≥ 24 (h) | 105 | 4.2 ± 2.0 | 2.29 ± 0.75 |
| Sanchez-Otero et al. [70] | Specific trained (n = 12) | Running | 0.14 (m s$^{-1}$ min$^{-1}$) | 0.14 (m s$^{-1}$ min$^{-1}$) | 23.9 ± 2.1 | 59.4 ± 5.1 | 15 (min) | ~ 105.3 | 3.0 ± 0.6 | 56.2 ± 4.7* |
| Midgley et al. [77] | Specific trained (n = 16) | Running | 0.28/0.14 (m s$^{-1}$ min$^{-1}$) | 0.28/0.14 (m s$^{-1}$ min$^{-1}$) | 11.6 ± 1.9 | 4.04 ± 0.46 | 10 (min) | ~ 102–103 | 2.8 ± 0.6 | 3.99 ± 0.45 |
| Scharhag-Rosenberger et al. [79] | Non-specific trained (n = 40) | Running | 0.19 (m s$^{-1}$ min$^{-1}$) | 0.19 (m s$^{-1}$ min$^{-1}$) | 15.0 ± 2.0 | 3.82 ± 0.99 | ≥ 24 (h) | 110 | 2.7 ± 0.6 | 3.75 ± 1.0 |
| Nolan et al. [90] | Non-specific trained (n = 12) | Running | 1 (% min$^{-1}$) | 1 (% min$^{-1}$) | 11.0 ± 1.8 | 56.9 ± 9.6 | 20 (min) | 105 | 2.6 ± 0.6 | 57.2 ± 9.0 |
| Nolan et al. [90] | Non-specific trained (n = 12) | Running | 1 (% min$^{-1}$) | 1 (% min$^{-1}$) | 11.4 ± 2.4 | 56.2 ± 9.0 | 60 (min) | 105 | 2.4 ± 0.7 | 56.2 ± 9.1 |
| Scharhag-Rosenberger et al. [79] | Non-specific trained (n = 40) | Running | 0.19 (m s$^{-1}$ min$^{-1}$) | 0.19 (m s$^{-1}$ min$^{-1}$) | 15.0 ± 2.0 | 3.82 ± 0.99 | 10 (min) | 110 | 2.1 ± 0.4 | 3.72 ± 0.99 |
| Astorino et al. [88] | Non-specific trained (n = 30) | Cycling | 23–29 (W min$^{-1}$) | 23–29 (W min$^{-1}$) | 7.7 ± 1.48 | 2.90 ± 0.60 | 10 (min) | ~ 106–109 | 2.1 ± 0.9 | 2.90 ± 0.60 |
| Astorino et al. [87] | Untrained (n = 9) | Cycling | 15–30 (W min$^{-1}$) | 15–30 (W min$^{-1}$) | 10.9 ± 1.26 | 2.67 ± 0.60 | 1–1.5 (h) | 115 | 2.0 ± 0.4 | 2.75 ± 0.76 |
| Keiller et al. [76] | Non-specific trained (n = 11) | Running | 1 (% min$^{-1}$) | 1 (% min$^{-1}$) | 8.8 ± 1.4 | 52.7 ± 5.64 | 6 (min) | ~ 110–112 | 1.9 ± 0.4 | 49.7 ± 4.30* |
| Astorino et al. [89] | Non-specific trained (n = 79) | Cycling | 6.7 (W min$^{-1}$) | 6.7 (W min$^{-1}$) | 25.6 ± 3.6 | 2.82 ± 0.62 | 10 (min) | 110 | 1.9 ± 0.4 | 2.78 ± 0.59 |
| McGawley et al. [68] | Specific trained (n = 10) | Running | 1 (% min$^{-1}$) | 1 (% min$^{-1}$) | 8.0 ± 0.9 | 59.2 ± 6.8 | 9 (min) | 105 | 1.8 ± 0.3 | 58.1 ± 6.7* |
| Nolan et al. [90] | Non-specific trained (n = 12) | Running | 1 (% min$^{-1}$) | 1 (% min$^{-1}$) | 11.2 ± 1.2 | 57.5 ± 9.2 | 20 (min) | 115 | 1.8 ± 0.5 | 56.9 ± 9.6 |
| Midgley et al. [37] | Specific trained (n = 10) | Cycling | 30 (W min$^{-1}$) | 30 (W min$^{-1}$) | 10.7 ± 1.3 | 4.05 ± 0.47 | 10 (min) | ~ 108 | ~ 1.7 | 3.96 ± 0.38 |
| Murias et al. [91] | Untrained (n = 45) | Cycling | 15–25 (W min$^{-1}$) | 15–25 (W min$^{-1}$) | ~ 10.5 | 40.6 ± 11.4 | 5 (min) | 105 | 1.7 ± 0.4 | 40.9 ± 10.9 |
| Astorino et al. [89] | Non-specific trained (n = 30) | Cycling | 20–40 (W min$^{-1}$) | 20–40 (W min$^{-1}$) | 10.1 ± 1.3 | 3.35 ± 1.01 | 8 (min) | 105 | 1.7 ± 0.4 | 3.32 ± 1.00 |
| Nolan et al. [90] | Non-specific trained (n = 12) | Running | 1 (% min$^{-1}$) | 1 (% min$^{-1}$) | 11.2 ± 1.5 | 57.1 ± 8.4 | 60 (min) | 115 | 1.7 ± 0.7 | 56.0 ± 9.3 |
sufficient durations, seem to be useful for the diagnosis of VO\textsubscript{2peak}, but not for the diagnosis of a VO\textsubscript{2pl} and VO\textsubscript{2max}.

### 3 Physiological Determinants

As described in the previous sections it seems to be very likely that the VO\textsubscript{2pl} is not simply a calculation artefact, as suggested by Beltrami et al. [32] and others [20, 24, 31, 33, 47]. However, there is also no convincing evidence for the assumption [11, 14, 34–36] that the absence of a VO\textsubscript{2pl} is mostly caused in inappropriate exercise modes and protocols or data analysing approaches. Thus, irrespective of the exercise modes and protocols approximately 20–60% of the participants show a plateau at VO\textsubscript{2max} when probably valid plateau definitions have been used [18, 19, 42, 49–52]. This raises the question of why some participants demonstrate a VO\textsubscript{2pl} and others do not.

#### 3.1 Why Does a VO\textsubscript{2}-Plateau Occur?

Hill and Lupton [3] interpreted the flattening of the VO\textsubscript{2}–running speed relationship at high running speeds as an indication that “the heart, lungs, circulation, and the diffusion of oxygen to the active muscle fibres have attained their maximum activity.” In a subsequent publication they identified the capacity of the cardiovascular and respiratory systems to transport O\textsubscript{2} as the main limiting factor of VO\textsubscript{2max} [1, 21]. The assumption that the VO\textsubscript{2pl} reflects the upper limit of the cardiovascular and respiratory system to transport O\textsubscript{2} to the muscles is supported by more recent studies [60, 100–102]. Brink-Elfegoun et al. [102] compared VO\textsubscript{2} and cardiac output (CO) using the Fick-principle during a constant load bout which barely elicits VO\textsubscript{2max} and during a 10–15% higher constant work rate. Despite the difference in work rate and sufficient duration of both constant load bouts, they found no differences in VO\textsubscript{2max} and CO between the higher and lower work rate bout. This indicates that the flattening of the VO\textsubscript{2}–work rate relationship at VO\textsubscript{2max} is caused by the inability of the heart to further increase the rate of O\textsubscript{2} delivery to the muscles. The finding of Brink-Elfegoun et al. [102] is supported by an older study from Miyamoto et al. [100] that described a close correlation between the work rates at which CO and VO\textsubscript{2} start to level off.

Harms et al. [101] showed that in female participants with a considerable arterial O\textsubscript{2} desaturation (\(\leq 92\%\)) at VO\textsubscript{2max} an increase of the O\textsubscript{2} partial pressure by breathing hyperoxic air (26% O\textsubscript{2} and 74% N\textsubscript{2}) leads to an increase of VO\textsubscript{2max} and removal of a VO\textsubscript{2pl} which was present at normoxic conditions. In contrast, participants without or only a slightly arterial O\textsubscript{2} desaturation (\(\geq 93\%\)) at VO\textsubscript{2max} showed no increase in VO\textsubscript{2max} when hyperoxic air was applied. Furthermore, they showed a VO\textsubscript{2pl} in both conditions [101].

| Authors (year) | Subjects (n) | Exercise mode | Incremental test | Verification test |
|---------------|--------------|---------------|-----------------|-------------------|
|               |              |               | Incremental rate | TTE (min) | VO\textsubscript{2max} (L min\(^{-1}\) / mL min\(^{-1}\) kg\(^{-1}\)) | Recovery duration | Intensity (%\textsubscript{Pmax}) | TTE (min) | VO\textsubscript{2max} (L min\(^{-1}\) / mL min\(^{-1}\) kg\(^{-1}\)) |
| Sansum et al. [69] | Untrained children (n = 128) | Cycling | 10–30 (W min\(^{-1}\)) | 9.5 ± 2.1 (m) 8.1 ± 2.0 (f) | 2.48 ± 0.73 (m) 1.96 ± 0.31 (f) | ~ 30 (min) | 105/110 | 1.6 ± 0.42 (m) 1.7 ± 0.33 (f) | 2.36 ± 0.72 (m)* 1.89 ± 0.34 (f)* |
| Rossiter et al. [92] | Non-specific trained (n = 7) | Cycling | 20 (W min\(^{-1}\)) | ~ 16.1 | 4.33 ± 0.52 5 (min) | 105 | 1.5 ± 0.3 | 4.30 ± 0.51 |
| Midgley et al. [37] | Specific trained (n = 10) | Running | 0.28 (m s\(^{-1}\) min\(^{-1}\)) | 11.4 ± 0.8 | 3.86 ± 0.39 10 (min) | ~ 105–107 | 1.5 | 3.92 ± 0.46 |
| Barker et al. [93] | Untrained children (n = 13) | Cycling | 10 (W min\(^{-1}\)) | 10.9 ± 1.5 | 1.69 ± 0.28 15 (min) | 105 | 1.5 ± 0.4 | 1.62 ± 0.31 |
| Sedgeman et al. [94] | Non-specific trained (n = 13) | Cycling | Not mentioned | 11.0 ± 1.7 | 50.1 ± 6.8 3 (min) | 105 | 1.3 ± 0.4 | 49.1 ± 6.7 |

Note that the bold indicates the studies with a sufficient duration of the verification bout (untrained > 3.5 min, non-specifically trained > 3.0 min, and specific endurance-trained > 2.0 min). TTE time to exhaustion, VO\textsubscript{2max} maximum oxygen uptake, m male subgroup, f female subgroup. *Significant difference between verification and incremental test VO\textsubscript{2max} Studies that reported separate findings of subgroups, e.g. because they performed in one subgroup treadmill tests and in the other one cycling tests, were presented for each subgroup separately.
Based on these findings it seems to be very likely that the VO₂pl is caused by the achievement of the highest rate at which the body can transport (and utilise) VO₂. This means that a real VO₂pl is caused by the same mechanisms that limit VO₂max, which are still controversially discussed [10, 22, 24, 33, 103–107]. However, for the occurrence of a VO₂pl it does not matter if VO₂max is limited by the diffusion of O₂ in the lung, the transport of O₂ by the cardiovascular system or the utilisation of O₂ in the muscles [103]. The crucial factor for the VO₂pl occurrence is that the mode-specific maximal rate of oxidative ATP generation (i.e. VO₂max) is reached before exercise is terminated due to fatigue or low levels of pain tolerance [49, 50, 103, 108, 109]. As described in Sect. 2.1, VO₂max must be sustained for rather large increases in work rate (~40–50 W) to enable VO₂pl diagnoses with sufficient certainty [42]. This raises the question of what kind of ability enables some participants to tolerate further increases in work rate or speed after VO₂max has been achieved.

3.2 Motivation and Pain Tolerance

Exercise in the heavy- and severe-intensity domain goes along with the perception of discomfort and pain. Therefore, the achievement of a VO₂pl has been ascribed to a high-level motivation and pain tolerance [2, 23, 35–37, 103, 109]. In fact, some studies (two of them with a probably valid VO₂pl definition [50, 51]) reported that the occurrence of a VO₂pl is accompanied by higher secondary exhaustion criteria, such as HRmax, RERmax and BLCmax [50, 51, 109]. However, most other studies (three of them with a probably valid VO₂pl definition [18, 19, 49]) reported no significant differences in secondary exhaustion criteria between plateauing and non-plateauing participants [15, 18, 19, 49, 73, 110–112]. This indicates that a high level of motivation and pain tolerance is only a necessary but not a sufficient requirement for the achievement of a VO₂pl. The latter conclusion is supported by a study of Doherty et al. [12] that found rather low VO₂pl incidences in elite runners, despite this cohort being accustomed to the discomfort and pain which goes along with high intensive exercise.

3.3 Anthropometric, Age and Gender Determinants

It is widely believed that children are less likely to demonstrate a VO₂pl compared to adults [2, 69, 73, 93, 112, 113]. This assumption is based on studies [58, 69, 73, 110], which tried to measure VO₂max in children and found considerable lower VO₂pl incidences than the classical study of Taylor et al. [7]. As described in Sect. 2.3.2, it is very likely, that the extremely high VO₂pl incidence in the study of Taylor et al. [7] is caused by the specific approach of this study, which results in a high rate of false-positive VO₂pl diagnoses [5, 52]. Compared to studies with a probably valid methodological approach [18, 19, 42, 49–52] the reported VO₂pl incidences in children are mostly in the range of adult participants (i.e. 20–60%). However, to the best of our knowledge, there is not a single study that directly compared the VO₂pl incidences of children and adults. Studies comparing young participants with and without a VO₂pl found no differences in age [16, 114, 115]. According to Edvardsen et al. [17] there was no effect of age on the VO₂pl incidence in a large cohort of more than 850 adult participants. Importantly, three of these studies used either an uncommon VO₂pl definition (no increase in VO₂ despite an increase in VE) [17] or fixed cut-offs without reporting the corresponding sampling intervals [16, 114]. As a result, it is unclear whether the findings of these studies are valid.

Beside of age, several studies observed the effect of height, body mass, body mass index, and sex on the probability of the VO₂pl occurrence. Only one study, which used a probably inappropriate VO₂pl definition, described a significantly higher VO₂pl incidence in female compared to male adults [116]. All other studies found no systematic effects of height, body mass, body mass index, and sex on the VO₂pl incidence [15–17, 69, 110, 112, 114, 115, 117]. Despite the validity of the VO₂pl definitions of some of these studies is unclear [16, 110, 112, 114, 116, 117], the large amount of nearly consistent evidence indicates that the VO₂pl occurrence is likely not affected by sex and anthropometric measurements. In contrast, the present evidence is insufficient to assess whether the VO₂pl incidence differs between children and adults or depends on age.

3.4 Aerobic Fitness and Endurance Training

It is partly believed that (endurance) trained athletes are more likely to show a plateau at VO₂max [2, 35, 36, 58, 103, 118]. This is based on the assumption that athletes are accustomed to sustained high-intensity exercise [103]. However, there is only one single study that compared the VO₂pl incidences between endurance-trained and untrained participants [51]. This study reported a significantly higher VO₂pl incidence in world-class cyclists (47%) compared to healthy sedentary participants (24%). In contrast, another study on elite runners found rather low VO₂pl incidences in incremental exercise tests that were performed on a treadmill [12]. Furthermore, markers of the VO₂pl occurrence are not correlated with VO₂max/VO₂peak in most studies observing the VO₂pl incidence in heterogeneous cohorts [13, 15, 16, 19, 43, 110]. This is supported by studies that found no differences in VO₂max/VO₂peak between participants with and without a VO₂pl [49, 50]. Furthermore, Gordon et al. [45] found a significantly lower VO₂max but equal VO₂pl incidences after a blood donation compared to the pre-blood donation test. Despite some of these studies used VO₂pl definitions with unclear or questionable validity [12, 15, 16, 45, 110].
this indicates that the VO_{2pl} occurrence is likely independent of aerobic fitness. Further studies are needed to verify whether endurance-trained participants that are accustomed to high-intensity exercise have higher VO_{2pl} incidences than untrained participants.

3.5 Anaerobic Power and Capacity

Since energy demand above VO_{2max} must be matched by anaerobic adenosine triphosphate (ATP) generation, the occurrence of a VO_{2pl} has been ascribed to differences in anaerobic power and capacity [17, 111, 112]. Thus, several studies observed the effect of anaerobic measurements on the VO_{2pl} occurrence [43, 108, 115, 119–122]. Most of them checked whether the occurrence of a VO_{2pl} depends on measurements of anaerobic power or force, like jumping height, running sprint performance, isokinetic knee extension/flexion or peak power of a 30 s Wingate-test [43, 115, 119, 120, 122]. However, it seems to be very unlikely that the maximum work rate of an incremental exercise test is limited by anaerobic power. Thus, peak-power of a 30 s Wingate-test or a 6 s sprint-test is 2–4 times higher than maximum work rate of an incremental test, which is used to measure VO_{2max} [123, 124]. Consequently, it is not surprising that none of these studies could find a systematic difference between plateauing and non-plateauing participants [43, 115, 119, 120, 122]. The ability to sustain VO_{2max} for long durations is much more likely affected by anaerobic capacity than anaerobic power, as shown for constant load exercise [125–127].

Based on these findings Gordon et al. [108] firstly checked whether the VO_{2pl} occurrence in incremental ramp tests depends on anaerobic capacity. They found a significant negative relationship between the maximal accumulated oxygen deficit and the increase in VO_{2} during the final minute of the ramp test, which indicates that participants with a higher anaerobic capacity are more likely to show a plateau at VO_{2max}. As shown in Fig. 5, this is probably caused by the fact that during incremental exercise a continuous accumulation of VO_{2}-deficit occurs due to a lagging behaviour of VO_{2} [49, 128, 129]. The VO_{2}-deficit must be matched by anaerobic energy contribution. This leads to the accumulation of muscle metabolites and, therefore, to exercise termination when a maximum tolerable concentration of fatigue-inducing metabolites has been achieved [49, 128–133]. At a given VO_{2}-deficit accumulation and VO_{2max}, a higher anaerobic capacity should enable a higher maximum work rate in an incremental exercise test. Participants with a higher anaerobic capacity are therefore able to tolerate a further increase in work rate or speed after VO_{2max} has been achieved, which may result in the occurrence of a VO_{2pl}, as shown in Fig. 5. This is supported by a recent study of Keiller and Gordon [134], which showed that the occurrence of the VO_{2pl} is associated with anaerobic alleles.

However, both studies [108, 134] used rather short sampling intervals (60 s or 60 breaths) combined with fixed cutoffs, such that the findings may be affected by elevated risks of false-positive and/or false-negative VO_{2pl} diagnoses. In addition, a more recent study by Silva et al. [121] did not find a correlation between the maximal accumulated oxygen deficit and measurements of the VO_{2pl} occurrence. This indicates that the VO_{2pl} occurrence is not solely affected by anaerobic capacity.

3.6 Oxygen Uptake Kinetics

The ability to sustain high-intensity exercise for a long duration is additionally determined by VO_{2}-kinetics [135, 136]. VO_{2}-kinetics is defined as the rate at which aerobic ATP generation adjusts to a change of exercise intensity [82]. The faster the VO_{2}-kinetics the lower the VO_{2}-deficit and the related accumulation of anaerobic metabolites at the beginning of exercise [136–139]. Consequently, faster VO_{2}-kinetics lowers the VO_{2}-deficit and spares anaerobic capacity, which leads to higher exercise tolerance and time to exhaustion during constant or intermittent exercise [140–143]. Furthermore, participants with faster VO_{2}-kinetics achieve their VO_{2max} earlier compared to counterparts with a slower VO_{2}-kinetics [49, 81].

The effect of VO_{2}-kinetics on the VO_{2pl} occurrence has been described recently [49]. This study demonstrated that participants with a VO_{2pl} have faster ramp- and square-wave VO_{2}-kinetics. As schematically shown in Fig. 6, the faster VO_{2}-kinetics leads to a lower VO_{2}-deficit accumulation in the submaximal intensity domain and results in an earlier achievement of VO_{2max}. As a result, participants with a faster VO_{2}-kinetics are able to sustain their VO_{2max} for longer durations, which increases the chance of a VO_{2pl} occurrence at the end of a ramp test [49].

Since VO_{2}-kinetics is speeded by a bout of priming exercise in the heavy/severe-intensity domain [85, 86], this finding seems to be supported by a study showing an increase in the VO_{2pl} incidence from 50 to 100% in ramp tests performed in a primed state compared to not primed ramp tests [40]. However, the authors used an inappropriate cut-off to check for the occurrence of a VO_{2pl}, which likely led to a high rate of false-positive VO_{2pl} diagnoses [42, 144]. Furthermore, a subsequent study did not find higher VO_{2pl} incidences in primed compared to not primed ramp tests [144]. Since the latter study found no speeding of ramp test VO_{2}-kinetics as well, this finding did neither support nor disprove a potential effect of VO_{2}-kinetics on the incidence of a VO_{2pl} [144].
In conclusion, evidence from a comparison of participants with and without a $\dot{V}O_2$pl suggests that $\dot{V}O_2$-kinetics is a determinant of the $\dot{V}O_2$-plateau occurrence. However, the cause–effect relationship needs to be proven by experimental research designs, which for example leads to speeding of $\dot{V}O_2$ ramp test kinetics, like priming exercise [145] or dietary nitrate supplementation [146].

### 3.7 Accumulation of Anaerobic Metabolites

Another factor that has been considered as a major determinant of the ability to sustain high intensity exercise is the relation of the work rate at which lactate begins to accumulate and the work rate that elicits $V_{O2max}$ [147–149]. This assumption is supported by significant negative relationships.
between lactate/ventilation thresholds expressed in %\(\dot{V}O_2\)max and time to exhaustion at the minimum running velocity that elicits \(\dot{V}O_2\)max [126, 149, 150].

Lacour et al. [50] analysed blood lactate and \(\dot{V}O_2\) values of stepwise incremental tests of 94 elite oarsmen, retrospectively. They found significantly lower blood lactate values in the submaximal intensity domain in the \(\dot{V}O_2\)pl group. As a result, the 4 mmol L\(^{-1}\) lactate threshold occurred at a significantly higher percentage of \(\dot{V}O_2\)max in the plateau compared to the non-plateau group. Furthermore, the %\(\dot{V}O_2\)max at the lactate threshold was closely negative correlated with the blood lactate concentration at the work rate step that firstly elicits \(\dot{V}O_2\)max. This seems to indicate that the oarsmen with a \(\dot{V}O_2\)pl are able to spare their anaerobic capacity to a greater extent because of a higher anaerobic threshold expressed in %\(\dot{V}O_2\)max.
compared to the non-plateauing oarsmen. The sparing of anaerobic capacity seems to enable the VO₂_{2pl} group to tolerate a larger increase in work rate after VO₂_{2max} has been achieved, which leads to the occurrence of a VO₂_{2pl} [50].

### 3.8 Central Governor

Based on low VO₂_{2pl} incidences Noakes [20, 24, 33] questioned the limitation of VO₂_{2max} and exercise performance by the provision of energy through aerobic and anaerobic ATP generation. Instead, he proposed a model in which the brain (central governor) regulates the recruitment of motor units to prevent damage of organs due to O₂ deficiency or loss of homeostasis [151]. The extensive but seemingly inconclusive discussion about the validity of the central governor model goes beyond the scope of the present review. The interested reader is referred to a substantial, however, clearly not complete body of related publications [20–22, 24, 33, 35, 152, 153].

Nevertheless, Noakes [20, 24, 33] stated repeatedly that in the absence of a VO₂_{2pl} termination of incremental exercise cannot be explained by the accumulation of anaerobic metabolites (muscle anaerobiosis). He strongly expressed the critique that since more than 20 years traditional concepts of exercise physiology failed to answer the following crucial question: “What causes the termination of exercise when the ‘true’ VO₂_{2max} is achieved without the ‘plateau phenomenon’?” [24]. Based on this supposedly unanswered question he concluded that “… the absence of the ‘plateau phenomenon’ in a majority of VO₂_{2max} tests can logically be interpreted in only one way: that factors other than a limiting cardiac output and the development of skeletal muscle anaerobiosis must cause the termination of exercise in the majority of VO₂_{2max} tests” [33]. This conclusion is based on the assumption that “the absence of a plateau indicates adequate muscle oxygenation during maximal exercise” [33], which would exclude any kind of oxygen deficit accumulation and also net anaerobic energy contribution, as long as no VO₂_{2pl} occurs.

However, the assumption that the absence of a VO₂_{2pl} indicates adequate muscle oxygenation per se and no need for muscular anaerobic energy provision reflect a misinterpretation of energy metabolism. As described in the previous sections there is a continuous accumulation of O₂-deficit from the beginning of an incremental exercise test due to a lagging behaviour of VO₂ [49, 128, 129, 132]. Furthermore, at work rates above maximal lactate steady-state or critical power metabolic energy is partly provided by anaerobic ATP-generation, which leads to the accumulation of anaerobic metabolites and muscular fatigue [130, 131, 146, 154–156]. Therefore, the termination of incremental exercise despite the absence of a VO₂_{2pl} can be well explained by VO₂-deficit accumulation, anaerobic energy contribution and the resulting accumulation of anaerobic metabolites [49, 50, 108, 128, 129, 148]. There is no need for a central governor to explain the presence or absence of a VO₂_{2pl}.

### 4 Consequences for the Diagnosis of VO₂_{2max}

As described in the introduction section one can be certain that VO₂_{2max} has been reached only if VO₂ remains more or less constant despite an increase in work rate. Thus, the diagnosis of VO₂_{2max} requires—per definition—a VO₂_{2pl}. However, inconsistencies in methodology and data processing in previous studies led to a lot of confusion about the correct diagnosis of the VO₂_{2pl}. When appropriate VO₂_{2pl} definitions and methods have been used the incidence of the VO₂_{2pl} during incremental exercise tests is usually less than 60% (see Sect. 2.1). Furthermore, there is no convincing evidence for the assumptions that a VO₂_{2pl} occurs in most participants when an incremental test is supplemented by a supramaximal verification bout or when a classical discontinuous exercise test is performed (see Sects. 2.3.2 and 2.3.3). This means that based on the present evidence VO₂_{2max} cannot be diagnosed in a considerable fraction of participants. Since this applies to often more than 50% of the participants it seems to be unreasonable to exclude all participants without a VO₂_{2pl}. As a consequence, we have to accept the diagnosis of VO_peak at least until there is more substantial evidence that most participants achieve a VO₂_{2pl} by applying a specific methodological approach.

Nevertheless, even if VO₂_{2max} cannot be diagnosed, it is important that VO_peak equals or is at least close to VO₂_{2max}. Otherwise, the efficiency of training interventions cannot be evaluated with sufficient certainty [11, 19]. Furthermore, a VO_peak which is considerably less than VO₂_{2max} may lead to fatal misdiagnoses since VO_peak is an important marker in clinical sports medicine and cardiology [157, 158]. The most common strategy to ensure that VO_peak is close to VO₂_{2max} is to apply secondary exhaustion criteria, like maximum blood lactate concentration, maximum heart rate, maximum respiratory exchange ratio or rating of or maximum rating of perceived exertion [2, 23, 25]. They can be used to reduce the magnitude of a potential underestimation of VO₂_{2max}, as recently described by Knaier et al. [18] and Wagner et al. [19]. However, the values of these criteria vary considerably between participants [37, 75] and are affected by the exercise protocol used [52, 62, 99]. Therefore, even if rather high and age-adjusted secondary exhaustion criteria are used an underestimation of VO₂_{2max} cannot be excluded [98]. Another strategy is to perform submaximal verification bouts (e.g. 95% of maximum work rate) after a common incremental test. As recently demonstrated this approach may be useful to identify submaximal VO_peak values of incremental tests,
especially in sedentary and clinical populations [98, 99], but also in recreationally active participants [42]. However, as described in Sect. 2.3.3 submaximal verification bouts do not allow for the diagnosis of a \( \dot{V}O_2 \) and corresponding \( \dot{V}O_2 \).

Based on this and the previous sections we recommend the following approach: For most studies in which \( \dot{V}O_2 / \dot{V}O_2 \) serves as a descriptive variable (like, age or sex) only, it seems to be sufficient to verify \( \dot{V}O_2 \) by the use of adequate and age-adjusted secondary exhaustion criteria [18, 19]. For all studies in which \( \dot{V}O_2 / \dot{V}O_2 \) is a main outcome, the \( \dot{V}O_2 \) profile should be checked for the occurrence of a \( \dot{V}O_2 \) by setting the cut-off at 50% of the increase in the submaximal intensity domain and using rather large sampling intervals (see Sect. 2.1). If no \( \dot{V}O_2 \) occurs a (submaximal) verification test should be performed, which allows for a sufficient duration of the verification bout (see Sect. 2.3.3) to ensure that \( \dot{V}O_2 \) is close as possible to \( \dot{V}O_2 \). This seems to be especially recommendable in sedentary and clinical populations [98, 99] In contrast, for highly endurance-trained participants this approach seems to be an insufficient because they do not achieve \( \dot{V}O_2 \) during short-lasting constant load tests [95–97].

## 5 Conclusions

A substantial fraction of studies used inappropriate \( \dot{V}O_2 \) definitions and approaches such that the validity of their findings is limited. As a result, it is unclear whether the \( \dot{V}O_2 \) incidence depends on exercise mode or incremental rate (see Sects. 2.2 and 2.3.1). Furthermore, there is no convincing evidence for the assumption that there is a higher \( \dot{V}O_2 \) incidence in classically discontinuous tests compared to continuous incremental tests (see Sect. 2.3.2). Since most studies that evaluated supramaximal verification phases, reported insufficient durations of the verification bouts, there is also rather limited evidence for an advantage of supramaximal verification tests approaches over single incremental test procedures in the determination of a \( \dot{V}O_2 \) and \( \dot{V}O_2 \) (see Sect. 2.3.3). Preliminary evidence suggests that the occurrence of a \( \dot{V}O_2 \) is determined by physiological factors like anaerobic capacity, \( \dot{V}O_2 \)-kinetics and accumulation of metabolites in the submaximal intensity domain.

Subsequent studies should take more attention to the use of valid \( \dot{V}O_2 \) definitions to avoid that their findings are biased by a high risk of false \( \dot{V}O_2 \) diagnoses. Therefore, cut-offs should be set at 50% of the corresponding increase in \( \dot{V}O_2 \) in the submaximal intensity domain and rather large sampling intervals should be used to enable an equal risk of false-positive and false-negative \( \dot{V}O_2 \) diagnoses as well as to account for the ventilation-induced variability of \( \dot{V}O_2 \) (see Sect. 2.1). If verification bouts are used to verify the achievement of \( \dot{V}O_2 / \dot{V}O_2 \), it should be ensured that they can be sustained for sufficient durations to enable \( \dot{V}O_2 \) to rise to the maximum value.

## Declarations

**Funding** Open Access funding enabled and organized by Projekt DEAL. Raphael Knaier was funded by the Swiss National Science Foundation (Grant P2BSP3_191755).

**Conflict of interest** Max Niemeyer, Raphael Knaier and Ralph Beneke declare that they have no conflict of interest or competing interests.

**Author Contribution Statement** MN: wrote the manuscript and created all figures and tables. RK: wrote and critically revised the manuscript. RB: critically revised the manuscript. All authors read and approved the manuscript.

**Open Access** This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article’s Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article’s Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit http://creativecommons.org/licenses/by/4.0/.

## References

1. Bassett DR. Scientific contributions of A. V. Hill: exercise physiology pioneer. J Appl Physiol. 2002;93(5):1567–82.
2. Howley ET, Bassett DR, Welch HG. Criteria for maximal oxygen uptake: review and commentary. Med Sci Sports Exerc. 1995;27(9):1292–301.
3. Hill AV, Lupton H. Muscular exercise, lactic acid, and the supply and utilization of oxygen. QJM. 1923;16(62):135–71.
4. Pokan R, Schwabeger G, Hofmann P, Eber B, Toplak H, Gasser R, Frühwald FM, Pessenhofer H, Klein W. Effects of treadmill exercise protocol with constant and ascending grade on levelling-off \( \dot{O}_2 \) uptake and \( \dot{V}O_2 \). Int J Sports Med. 1995;16(4):238–42.
5. Glassford RG, Baycroft GH, Sedgwick AW, Macnab RB. Comparison of maximal oxygen uptake values determined by predicted and actual methods. J Appl Physiol. 1965;20(3):509–13.
6. Mitchell JH, Sproule BJ, Chapman CB. The physiological meaning of the maximal oxygen intake test. J Clin Invest. 1958;37(4):538–47.
7. Taylor HL, Buskirk E, Henschel A. Maximal oxygen intake as an objective measure of cardio-respiratory performance. J Appl Physiol. 1955;8(1):73–80.
8. Levine BD. \( \dot{V}O_{2\text{max}} \): what do we know, and what do we still need to know? J Physiol. 2008;586(1):25–34.
9. Ross R, Blair SN, Arena R, Church TS, Després J-P, Franklin BA, Haskell WL, Kaminsky LA, Levine BD, Lavie CJ, Myer J, Niebauer J, Sallis J, Sawada SS, Sui X, Wilsöf U. Importance of assessing cardiorespiratory fitness in clinical practice: a case
for fitness as a clinical vital sign: a scientific statement from the American heart association. Circulation. 2016;134(24):e653–99.

10. Bassett DR, Howley ET. Limiting factors for maximum oxygen uptake and determinants of endurance performance. Med Sci Sports Exerc. 2000;32(1):70–84.

11. Poole DC, Jones AM. Measurement of the maximum oxygen uptake $VO_{2\text{max}}$: $VO_{2\text{peak}}$ is no longer acceptable. J Appl Physiol. 2017;122(4):997–1002.

12. Doherty M, Nobbs L, Noakes TD. Low frequency of the “plateau phenomenon” during maximal exercise in elite British athletes. Eur J Appl Physiol. 2003;89(6):619–23.

13. Day JR, Rossiter HB, Coats EM, Skasick A, Whipp BJ. The maximally attainable VO2 during exercise in humans: the peak vs maximum issue. J Appl Physiol. 2003;95(5):1901–7.

14. Astorino TA. Alterations in VO2max and the VO2 plateau with manipulation of sampling interval. Clin Physiol Funct Imaging. 2009;29(1):60–7.

15. Wood RE, Hills AP, Hunter GR, King NA, Byrne NM. VO2max in overweight and obese adults: do they meet the threshold criteria? Med Sci Sports Exerc. 2010;42(3):470–7.

16. Peyer K, Pivarnik JM, Coe DP. The relationship among HRpeak, RERpeak, and VO2 peak during treadmill testing in girls. Res Q Exerc Sport. 2011;82(4):685–92.

17. Edvardsen E, Hem E, Anderssen SA. End criteria for reaching the VO2 max: plateaus might exist. J Appl Physiol. 1998;85(4):1391–9.

18. Knaier R, Niemeyer M, Wagner J, Infanger D, Hinrichs T, Klenk C, Frutig S, Cajoochen C, Schmidt-TRUCKSIS S A. Which cut-offs for secondary VO2max criteria are robust to diurnal variations? Med Sci Sports Exerc. 2019;51(5):1006–13.

19. Wagner J, Niemeyer M, Infanger D, Hinrichs T, Streese L, Hansen H, Nylander J, Schmidt-TRUCKSIS S A, Knaier R. New database cut-offs for maximal exercise criteria across the lifespan. Med Sci Sports Exerc. 2020;52(9):1915–23.

20. Noakes TD. Implications of exercise testing for prediction of athletic performance: a contemporary perspective. Med Sci Sports Exerc. 1988;20(4):319–30.

21. Bassett DR, Howley ET. Maximal oxygen uptake: “classical” versus “contemporary” viewpoints. Med Sci Sports Exerc. 1997;29(5):591–603.

22. Bergh U, Ekblom B, Astrand PO. Maximal oxygen uptake “classical” versus “contemporary” viewpoints. Med Sci Sports Exerc. 2000;32(1):85–8.

23. Midgley AW, McNaughton LR, Polman R, Marchant D. Criteria for determination of maximal oxygen uptake: a brief critique and recommendations for future research. Sports Med. 2007;37(12):1019–28.

24. Noakes TD. How did A. V. Hill understand the VO2max and the “plateau phenomenon”? Still no clarity? Br J Sport Med. 2008;42(7):574–80.

25. Robers GA, Dwyer D, Astorino T. Recommendations for improved data processing from expired gas analysis indirect calorimetry. Sports Med. 2010;40(2):95–111.

26. Smirmaul BPC, Bertucci DR, Teixeira IP. Is the VO2max that we measure really maximal? Front Physiol. 2013;4:203.

27. Beltz NM, Gibson AL, Janot JM, Kravitz L, Mermier CM, Daleck LC. Graded exercise testing protocols for the determination of VO2max: historical perspectives, progress, and future considerations. J Sports Med (Hindawi Publ Corp). 2016;2016:3968393.

28. Schaun GZ. The maximal oxygen uptake verification phase: a light at the end of the tunnel? Sports Med Open. 2017;3(1):44.

29. Pettit RW, Jammick NA. Commentary on “Measurement of the maximum oxygen uptake VO2max: VO2 peak is no longer acceptable.” J Appl Physiol. 2017;123(3):696.
49. Niemeyer M, Leithaeuser R, Beneke R. Oxygen uptake plateau occurrence depends on oxygen kinetics and oxygen deficit accumulation. Scand J Med Sci Sports. 2019;29(10):1466–72.
50. Lacour J-R, Messonnier L, Bourdin M. The leveling-off of oxygen uptake is related to blood lactate accumulation Retrospective study of elite rowers. Eur J Appl Physiol. 2007;101(2):241–7.
51. Lucia A, Rabadán M, Hoyos J, Hernández-Capilla M, Pérez M, San Juan AF, Earnest CP, Chicharro JL. Frequency of the VO2max plateau phenomenon in world-class cyclists. Int J Sports Med. 2006;27(12):984–92.
52. Duncan GE, Howley ET, Johnson BN. Applicability of VO2max criteria: discontinuous versus continuous protocols. Med Sci Sports Exerc. 1997;29(2):273–8.
53. Shephard RJ, Allen C, Benade AJ, Davies CT, Di Prampero PE, Hedman R, Merriman JE, Myhre K, Simmons R. The maximum oxygen intake. An international reference standard of cardiorespiratory fitness. Bull World Health Organ. 1968;38(5):757–64.
54. Astrand PO, Saltin B. Maximal oxygen uptake and heart rate in various types of muscular activity. J Appl Physiol. 1961;16:977–81.
55. Wyndham CH, Strydom NB, Leary WP, Williams CG. Studies of the maximum capacity of men for physical effort. I. A comparison of methods of assessing the maximum oxygen intake. Int J Angew Physiol. 1966;22(4):285–95.
56. Millet GP, Vleck VE, Bentley DJ. Physiological differences between cycling and running: lessons from triathletes. Sports Med. 2009;39(3):179–206.
57. Stromme SB, Ingjer F, Meen HD. Assessment of maximal aerobic power in specifically trained athletes. J Appl Physiol Respir Environ Exerc Physiol. 1977;42(6):833–7.
58. Cumming GR, Friesen W. Bicycle ergometer measurement of maximal oxygen uptake in children. Can J Physiol Pharmacol. 1967;45(6):937–46.
59. Hermansen L, Saltin B. Oxygen uptake during maximal treadmill and bicycle exercise. J Appl Physiol. 1969;26(1):31–7.
60. Hermansen L, Ekblom B, Saltin B. Cardiac output during submaximal and maximal treadmill and bicycle exercise. J Appl Physiol. 1970;29(1):82–6.
61. Brink-Elfegoun T, Holmberg H-C, Ekblom MN, Ekblom B. Neuromuscular and circulatory adaptation during combined arm and leg exercise with different maximal work loads. Eur J Appl Physiol. 2007;101(5):603–11.
62. Adami A, Sivieri A, Moia C, Perini R, Ferretti G. Effects of step duration in incremental ramp protocols on peak power and maximal oxygen consumption. Eur J Appl Physiol. 2013;113(10):2647–53.
63. Midgley AW, Bentley DJ, Luttikholt H, McNaughton LR, Millet GP. Challenging a dogma of exercise physiology: does an incremental exercise test for valid VO2max determination really need to last between 8 and 12 minutes? Sports Med. 2008;38(6):441–7.
64. Yoon B-K, Kravitz L, Robergs R. VO2max protocol duration, and the VO2 plateau. Med Sci Sports Exerc. 2007;39(7):1186–92.
65. Buchfuhrer MJ, Hansen JE, Robinson TE, Sue DY, Wasserman K, Whipp BJ. Optimizing the exercise protocol for cardiopulmonary assessment. J Appl Physiol Respir Environ Exerc Physiol. 1983;55(5):1558–64.
66. Astorino TA, Rietschel JC, Tam PA, Taylor K, Johnson SM, Freedman TP, Sakarya CE. Reinvestigation of optimal duration of VO2max testing. JEPonline. 2004;7:1–8.
67. Weston SB, Gray AB, Schneider DA, Gass GC. Effect of ramp slope on ventilation thresholds and VO2peak in male cyclists. Int J Sports Med. 2002;23(1):22–7.
68. McGawley K. The reliability and validity of a four-minute running trial in assessing VO2max and performance. Front Physiol. 2017;8:270.
69. Sansum KM, Weston ME, Bond B, Cockcroft DJ, O’Connor A, Tomlinson OW, Williams CA, Barker AR. Validity of the supramaximal test to verify maximal oxygen uptake in children and adolescents. Pediatr Exerc Sci. 2019;31(2):213–22.
70. Sánchez-Otero T, Iglesias-Soler E, Boullosa DA, Tuinil JL. Verification criteria for the determination of VO2max in the field. J Strength Cond Res. 2014;28(12):3544–51.
71. Binkhorst RA, van Leeuwen. A rapid method for the determination of aerobic capacity. Int J Angew Physiol. 1963;19:459–67.
72. Froelicher VF, Brammell H, Davis G, Nogueria I, Stewart A, Lancaster MC. A comparison of three maximal treadmill exercise protocols. J Appl Physiol. 1974;36(6):720–5.
73. Astrand PO. Experimental studies of physical working capacity in relation to sex and age. Copenhagen: Munksgaard Forlag; 1952.
74. Midgley AW, Carroll S. Emergence of the verification phase procedure for confirming ‘true’ VO2max. Scand J Med Sci Sports. 2009;19(3):313–22.
75. Poole DC, Wilkerson DP, Jones AM. Validity of criteria for establishing maximal O2 uptake during ramp exercise tests. Eur J Appl Physiol. 2008;102(4):403–10.
76. Keiffer D, Gordon D. Confirming maximal oxygen uptake: is hear rate the answer? Int J Sports Med. 2018;39(3):198–203.
77. Midgley AW, McNaughton LR, Carroll S. Verification phase as a useful tool in the determination of the maximal oxygen uptake of distance runners. Appl Physiol Nutr Metab. 2006;31(5):541–8.
78. Meer CM, Alexander RP, Mageean AL. Achievement of VO2max criteria during a continuous graded exercise test and a verification stage performed by college athletes. J Strength Cond Res. 2012;26(10):2648–54.
79. Scharhag-Rosenberger F, Carlsohn A, Cassel M, Mayer F, Scharhag J. How to test maximal oxygen uptake: a study on timing and testing procedure of a supramaximal verification test. Appl Physiol Nutr Metab. 2011;36(1):153–60.
80. Hill DW, Poole DC, Smith JC. The relationship between power and the time to achieve VO2max. Med Sci Sports Exerc. 2002;34(4):709–14.
81. Caputo F, Denadai BS. The highest intensity and the shortest duration permitting attainment of maximal oxygen uptake during cycling: effects of different methods and aerobic fitness level. Eur J Appl Physiol. 2008;103(1):47–57.
82. Poole DC, Jones AM. Oxygen uptake kinetics. Compr Physiol. 2012;2(2):933–96.
83. Koppo K, Bouckaert J, Jones AM. Effects of training status and exercise intensity on phase II VO2 kinetics. Med Sci Sports Exerc. 2004;36(2):225–32.
84. Caputo F, Mello MT, Denadai BS. Oxygen uptake kinetics and time to exhaustion in cycling and running: a comparison between trained and untrained subjects. Arch Physiol Biochem. 2003;111(5):461–6.
85. Jones AM, Koppo K, Burnley M. Effects of prior exercise on metabolic and gas exchange responses to exercise. Sports Med. 2003;33(13):949–71.
86. Bailey SJ, Vizhatalo A, Wilkerson DP, DiMenna FJ, Jones AM. Optimizing the “priming” effect: influence of prior exercise intensity and recovery duration on O2 uptake kinetics and severe-intensity exercise tolerance. J Appl Physiol (1985). 2009;107(6):1743–56.
87. Astorino TA, White AC, Dallec LC. Supramaximal testing to confirm attainment of VO2max in sedentary men and women. Int J Sports Med. 2009;30(4):279–84.
88. Astorino TA, White AC. Assessment of anaerobic power to verify VO2max attainment. Clin Physiol Funct Imaging. 2010;30(4):294–300.
93. Barker AR, Williams CA, Jones AM, Armstrong N. Establishing maximal oxygen uptake during a ramp incremental test is not confirmed by a verification phase. Front Physiol. 2018;9:143.

94. Rossiter HB, Kowalchuk JM, Whipp BJ. A test to establish maximal $\dot{V}O_2$ uptake during no plateau in the $\dot{O}_2$ uptake response to ramp incremental exercise. J Appl Physiol. 2006;100(3):764–70.

95. Draper SB, Wood DM, Fallowfield JL. The VO2 response to exhaustive severe-intensity running. Int J Sports Med. 2017;8:96.

96. Nolan PB, Beaven ML, Dalleck L. Comparison of intensities and mode. Eur J Appl Physiol. 2003;90(1–2):361–74.

97. Black MJ, Potter CR, Corbett J, Clark CCT, Draper SB. Maximal oxygen uptake is achieved in hypoxia but not normoxia during an exhaustive severe intensity run. Front Physiol. 2017;8:96.

98. Arad AD, Bishop K, Adimoolam D, Albu JB, DiMenna FJ. Severe-intensity constant-work-rate cycling indicates that ramp incremental cycling underestimates $\dot{V}O_2max$ in a heterogeneous cohort of sedentary individuals. PLoS ONE. 2020;15(7):e0235567.

99. Bowen TS, Cannon DT, Begg G, Baliga V, Witte KK, Ros-siter HB. A novel cardiopulmonary exercise test protocol and criterion to determine maximal oxygen uptake in chronic heart failure. J Appl Physiol (1985). 2012;113(3):451–8.

100. Miyamoto Y, Nizeki K, Abe H. Behavior of cardiac output during progressive exercise tests: a preliminary report. Ann Physiol Anthropol. 1992;11(3):225–30.

101. Harms CA, McClaran SR, Nickele GA, Pegelow DF, Niel-son WB, Dempsey JA. Effect of exercise-induced arterial $O_2$ desaturation on $\dot{V}O_2max$ in women. Med Sci Sports Exerc. 2000;32(6):1101–8.

102. Brink-Ellégou T, Kajser L, Gustafsson T, Ekblom B. Maximal oxygen uptake is not limited by a central nervous system govern-ner. J Appl Physiol. 2007;102(2):781–6.

103. Wagner PD. New ideas on limitations to $\dot{V}O_2max$. Exerc Sport Sci Rev. 2000;28(1):10–4.

104. Saltin B, Calbet JAL. Point: in health and in a normoxic environment, $\dot{V}O_2max$ is limited primarily by cardiac output and locomotor muscle blood flow. J Appl Physiol. 2006;100(2):744–5.

105. Wagner PD. Counterpoint: in health and in normoxic environment $\dot{V}O_2max$ is limited primarily by cardiac output and locomotor muscle blood flow. J Appl Physiol. 2006;100(2):745–7.

106. Lundby C, Montero D. Rebuttal from Carsten Lundby and David Montero. J Physiol. 2015;593(17):3765.

107. Wagner PD. CrossTalk proposal: Diffusion limitation of $O_2$ from microvessels into muscle does contribute to the limitation of $\dot{V}O_2max$. J Physiol. 2015;593(17):3757–8.

108. Gordon D, Hopkins S, King C, Keiller D, Barnes RJ. Incidence of the plateau at $\dot{V}O_2max$ is dependent on the anaerobic capacity. Int J Sports Med. 2011;32(1):1–6.

109. Sidney KH, Shephard RJ. Maximum and sub-maximum exercise tests in men and women in the seventh, eighth, and ninth decades of life. J Appl Physiol Respir Environ Exerc Physiol. 1977;43(2):280–7.

110. Armstrong N, Kirby BJ, McManus AM, Welsman JR. Aerobic fitness of prepubescent children. Ann Hum Biol. 1995;22(5):427–41.

111. Cunningham DA, van Waterschoot BM, Paterson DH, Lef-coe M, Sangal SP. Reliability and reproducibility of maximal oxygen uptake measurement in children. Med Sci Sports. 1977;9(2):104–8.

112. Mahon AD, Marsh ML. Ventilatory threshold and $\dot{V}O_2$ plateau at maximal exercise in 8- to 11-year-old children. Pediatr Exerc Sci. 1993;5(4):332–8.

113. Kropej V, Škof B, Milić R. The influence of aerobic and anaerobic characteristics of children of different age on achievement of $\dot{V}O_2plateau$. Biolog Sport. 2007;24:13–9.

114. Armstrong N, Welsman J, Winsley R. Is peak $\dot{V}O_2$ a maximal index of children’s aerobic fitness? Int J Sports Med. 1996;17(5):356–9.

115. Misquita NA, Davis DC, Dobrovolny CL, Ryan AS, Dennis KE, Nicklas BJ. Applicability of maximal oxygen consumption criteria in obese, postmenopausal women. J Womens Health Gend Based Med. 2001;10(9):879–85.

116. Krobe V, Škof B, Milić R. The influence of aerobic and anaerobic characteristics of children of different age on achievement of $\dot{V}O_2plateau$. Biolog Sport. 2005;22:67–79.

117. Armstrong N, Kirby BJ, McManus AM, Welsman JR. Maximum and sub-maximum exercise tests in men and women in the seventh, eighth, and ninth decades of life. J Appl Physiol Respir Environ Exerc Physiol. 1977;43(2):280–7.

118. Misquita NA, Davis DC, Dobrovolny CL, Ryan AS, Dennis KE, Nicklas BJ. Applicability of maximal oxygen consumption criteria in obese, postmenopausal women. J Womens Health Gend Based Med. 2001;10(9):879–85.

119. Rivera-Brown AM, Alvarez M, Rodríguez-Santana JR, Benetti PJ. Anaerobic power and achievement of $\dot{V}O_2max$ in prepubertal boys. Int J Sports Med. 2001;22(2):111–5.

120. Rowland TW, Cunningham LN. Oxygen uptake plateau during maximal treadmill exercise in children. Chest. 1992;101(2):485–9.

121. Silva RG, Silva-Cavalcante MD, Azevedo RDA, Lima-Silva AE, Bertuzzi R. The $\dot{V}O_2max$ plateau is not associated with the anaerobic capacity in physically active subjects. Rev bras educ fis esporte. 2016;30(4):857–64.

122. Duncan GE, Mahon AD, Howe CA, Del Corral P. Plateau in oxygen uptake at maximal exercise in male children. Pediatr Exerc Sci. 1996;8(1):77–86.

123. Bénéke R, Pollmann C, Bleif I, Leithäuser RM, Hütler M. How occurs in elite cyclists, kayakists and swimmers. Eur J Appl Physiol Occup Physiol. 1997;76(1):13–20.

124. Rowland TW, Cunningham LN. Oxygen uptake plateau during maximal treadmill exercise in children. Chest. 1992;101(2):485–9.

125. Billat V, Beaillot J, Jan J, Rochcongar P, Carre F. Gender effect on $\dot{V}O_2max$. J Physiol. 2015;593(17):3765.

126. McCoy AE, Bertuzzi R. The $\dot{V}O_2max$ plateau is dependent on the anaerobic capacity. Int J Sports Med. 2001;22(2):111–5.

127. Billat V, Beaillot J, Jan J, Rochcongar P, Carre F. Gender effect on the relationship of time limit at 100% $\dot{V}O_2max$ with other bioenergetic characteristics. Med Sci Sports Exerc. 1996;28(8):1049–55.

128. Hill DW, Rowell AL. Significance of time to exhaustion during exercise at the velocity associated with $\dot{V}O_2max$. Eur J Appl Physiol Occup Physiol. 1996;72(4):383–6.

129. Faina M, Billat V, Squadrone R, de Angelis M, Koralszttein JP, Dal Monte A. Anaerobic contribution to the time to exhaustion at the minimal exercise intensity at which maximal oxygen uptake occurs in elite cyclists, kayakists and swimmers. Eur J Appl Physiol Occup Physiol. 1997;76(1):13–20.

130. Pouilly J-P, Busso T. Accumulated oxygen deficit during ramp exercise. Int J Sports Med. 2008;29(1):16–20.

131. Scott CB, Bogdanffy GM. Aerobic and anaerobic energy expenditure during exhaustive ramp exercise. Int J Sports Med. 1998;19(4):277–80.
