Measurement of a True V̇O$_{2max}$ during a Ramp Incremental Test Is Not Confirmed by a Verification Phase

Juan M. Murias[1]*, Silvia Pogliaghi[2] and Donald H. Paterson[3]

[1] Faculty of Kinesiology, University of Calgary, Calgary, AB, Canada, [2] Department of Neurosciences, Biomedicine and Movement Sciences, University of Verona, Verona, Italy, [3] School of Kinesiology, The University of Western Ontario, London, ON, Canada

The accuracy of an exhaustive ramp incremental (RI) test to determine maximal oxygen uptake (V̇O$_{2max}$) was recently questioned and the utilization of a verification phase proposed as a gold standard. This study compared the oxygen uptake (V̇O$_2$) during a RI test to that obtained during a verification phase aimed to confirm attainment of V̇O$_{2max}$. Sixty-one healthy males [31 older (O) 65 ± 5 yrs; 30 younger (Y) 25 ± 4 yrs] performed a RI test (15–20 W/min for O and 25 W/min for Y). At the end of the RI test, a 5-min recovery period was followed by a verification phase of constant load cycling to fatigue at either 85% ($n = 16$) or 105% ($n = 45$) of the peak power output obtained from the RI test. The highest V̇O$_2$ after the RI test (39.8 ± 11.5 mL·kg^{-1}·min^{-1}) and the verification phase (40.1 ± 11.2 mL·kg^{-1}·min^{-1}) were not different ($p = 0.33$) and they were highly correlated ($r = 0.99$; $p < 0.01$). This response was not affected by age or intensity of the verification phase. The Bland-Altman analysis revealed a very small absolute bias (−0.25 mL·kg^{-1}·min^{-1}, not different from 0) and a precision of ±1.56 mL·kg^{-1}·min^{-1} between measures. This study indicated that a verification phase does not highlight an under-estimation of V̇O$_{2max}$ derived from a RI test, in a large and heterogeneous group of healthy younger and older men naïve to laboratory testing procedures. Moreover, only minor within-individual differences were observed between the maximal V̇O$_2$ elicited during the RI and the verification phase. Thus a verification phase does not add any validation of the determination of a V̇O$_{2max}$. Therefore, the recommendation that a verification phase should become a gold standard procedure, although initially appealing, is not supported by the experimental data.

Keywords: maximal oxygen uptake, oxidative metabolism, exercise testing, aerobic performance, aging

INTRODUCTION

Exercise physiologists have been interested in the measurements of oxygen uptake (V̇O$_2$) and maximal V̇O$_2$ (V̇O$_{2max}$) since early in the 20th century (Krogh and Lindhard, 1913; Hill and Lupton, 1923). Given the importance of V̇O$_{2max}$ as an integrative measure of different components of the cardiovascular and neuromuscular system to exercise, its measurement has been widespread not only as an indicator of human performance (Hoppeler and Weibel, 2000; di Prampero, 2003; Levine, 2008), but also as a marker of overall cardiovascular and cardiorespiratory function in different healthy and clinical populations

(Frontera et al., 1990; Borrelli et al., 2003; Pogliaghi et al., 2006; Levine, 2008; Paterson and Warburton, 2010; Jensen et al., 2016) and different environmental conditions (Bringard et al., 2010; Doria et al., 2011). Similarly, measures of V̇O$_{2max}$ have been used to determine the efficacy of different exercise training interventions aimed at improving physiological function both from performance as well as health perspectives (Pogliaghi et al., 2006; Murias et al., 2010a; Bruseghini et al., 2015). Considering the importance of V̇O$_{2max}$ for assessment of performance and health, adequate protocols capable of establishing a true maximal value are needed to provide confidence for evaluation and longitudinal follow-up.

Early studies measuring expired gases and volumes from a Douglas bag promoted the use of testing protocols that required prolonged steps of progressively increased intensity; such protocols aimed to establish a steady-state or sustainable response in each step, until critical intensities of exercise prevented further steps from being completed (ÅSTRAND, 1960; Glassford et al., 1965). Under those conditions, a work rate (or speed) of exercise that resulted in no further increase in V̇O$_2$ despite the increase in energy demand was established and this plateau response was considered a true V̇O$_{2max}$, as proposed by Taylor et al. (1955). With the development of fast-response breath-by-breath gas exchange and volume analyzers, researchers started to move away from these types of time-consuming protocols (Buchfuhrer et al., 1983), and step incremental and ramp incremental (RI) tests became widely used to measure the V̇O$_{2max}$ response during exercise to exhaustion. RI protocols allow for rapid determination of V̇O$_{2max}$ as well as other valuable indexes such as the exercise intensity thresholds; however, some researchers have questioned the accuracy of RI tests to consistently provide a true V̇O$_{2max}$ value as a plateau in V̇O$_2$ is not always (or seldom) observed (Rossiter et al., 2006; Poole et al., 2008; Poole and Jones, 2017). To circumvent this limitation, secondary criteria such as an increase in blood lactate concentration [La] above 8 mmol·L^{-1}, a heart rate response within 10 beats per minute (bpm) of the maximal predicted value, a respiratory exchange ratio (RER) higher than 1.10, and a rating of perceived exertion (RPE) >18, are commonly used to establish the attainment of a true V̇O$_{2max}$ response (ACSM, 2003). However, as indicated by Midgley et al. (2007) and Poole et al. (2008), secondary criteria do not seem valid for accurate determination of the attainment of V̇O$_{2max}$.

Another approach to establish the attainment of V̇O$_{2max}$ consists of executing a verification (or validation) phase subsequent to the RI test (Day et al., 2003; Midgley and Carroll, 2009). Although the exact origin of this model is not clearly established (Midgley and Carroll, 2009), Rossiter et al. (2006) proposed that, following a recovery period of five minutes subsequent to the RI test, a constant load exercise to exhaustion should be performed at either 95 or 105% of the peak power output (PO) obtained during the RI test. The use of a verification phase to confirm the attainment of a true maximum during a RI test in which a plateau in V̇O$_2$ may not be present, is based on the following assumptions:

(i) a RI test may fail to provide a true V̇O$_{2max}$; (ii) an upper limit to V̇O$_2$ exists and this value can only be underestimated but not overestimated; (iii) during a constant load exercise to the limit of tolerance in the very heavy/severe domain, V̇O$_2$ will project to V̇O$_{2max}$ within the time of task failure. Therefore, according to Rossiter et al. (2006), the lack of significant differences in the V̇O$_2$ response of the constant load verification phase, both at 95 and 105% of peak PO, compared to the highest V̇O$_2$ observed during the RI test, would provide the objective "plateau criterion" that often remains elusive during a RI test and confirm that a maximal value was obtained. Importantly, this protocol not only corroborated the presence of an upper limit of the V̇O$_2$ response, but it also indicated that the highest V̇O$_2$ values obtained during a RI were not different from those observed during the verification phase and likely reflected the achievement of a true V̇O$_{2max}$.

Recently, a review by Poole and Jones (2017) proposed the idea that the denomination of "peak" V̇O$_2$ (V̇O$_{2peak}$) is no longer acceptable and that validated V̇O$_{2max}$ results, as derived from a verification phase, should be presented in all future studies. An important point raised by Poole and Jones (2017) is that even though the RI test might yield a highly reproducible V̇O$_{2max}$ response in active or trained participants who are accustomed to pushing themselves to exhaustion, this may not be the case with less experienced, unmotivated, and/or clinical populations. Thus, even though a RI test might provide a trustable measure of V̇O$_{2max}$ in trained individuals, a verification phase should always be performed in any population unaccustomed to pushing to the upper limits of tolerance.

Although the verification phase appears appealing in providing a strategy to overcome the ongoing V̇O$_2$ max-peak debate, data from different studies do not convincingly support the usefulness or necessity of a verification phase either in the general healthy population or in specific clinical/frail/unfit populations. In sedentary (Astorino et al., 2009), recreationally active (Sedgeman et al., 2013; Nolan et al., 2014), overweight and obese (Wood et al., 2010; Sawyer et al., 2015) adults, as well as children (Barker et al., 2011), the V̇O$_2$ response during the verification phase was not higher than that observed during the RI test. This was also the case in a group of chronic heart failure patients (Bowen et al., 2012). These data could suggest that the verification phase provided a confirmation that a true V̇O$_{2max}$ was established in these subjects. However, these very same findings could just as plausibly indicate that the RI test *per-se* was an adequate protocol, yielding a true maximal response in most participants. Moreover, a recent study has shown that a verification phase resulted in a lower V̇O$_2$ value as compared to that observed during a graded incremental test (McGawley, 2017). Importantly, the possible usefulness of a verification phase has only been investigated in relatively small and homogeneous samples and (with the exception of McGawley (2017)) with a suboptimal statistical approach (i.e., only comparison of mean values). Additionally, healthy older adults, for whom V̇O$_{2max}$ is the strongest predictor of independent living (Paterson and Warburton, 2010) and for whom a verification phase would theoretically be needed for accurate V̇O$_{2max}$ determination (Poole and Jones, 2017), have never previously been evaluated. Establishing if a verification phase actually adds confidence in the

achievement of $\dot{V}O_{2max}$ in this population would contribute to further support or oppose the recommendation of such practice.

Poole and Jones (2017) recommended that the verification phase should be performed at intensities exceeding the peak PO obtained at the end of the RI test (e.g., 105% of peak PO). This is proposed in order to satisfy the plateau criterion (i.e., no further increase in $\dot{V}O_2$ despite the increase in PO) that is the foundation of the $\dot{V}O_{2max}$ concept. However, it should be noted that, due to the presence of the so called slow component of $\dot{V}O_2$, any constant load PO above critical power will result in the achievement of $\dot{V}O_{2max}$, provided that time to task failure is sufficiently prolonged (Poole and Jones, 2012). This well-known physiological phenomenon provides the rationale for using workloads that are not only above but also below maximal PO for the verification phase (Rossiter et al., 2006), with lower intensities possibly favoring compliance and reducing the risk associated to supramaximal exercise in older or clinical populations. Indeed, verification phases conducted at a submaximal PO yield $\dot{V}O_{2max}$ values not different from those obtained from RI tests (Rossiter et al., 2006; Sedgeman et al., 2013).

Thus, given the uncertainties related to the ability of establishing $\dot{V}O_{2max}$ during a RI test, and the proposal that a verification phase with constant load might confirm the RI $\dot{V}O_2$ or identify that the RI test was not maximal, the goals of this study were to: (1) determine whether a constant load verification phase yields a $\dot{V}O_2$ higher than the RI test; (2) examine the variation of the differences around the mean between the highest $\dot{V}O_2$ of the RI test and the verification phase; (3) determine the role of age group (i.e., younger vs. older adults) on the possible difference in the $\dot{V}O_2$ value between the RI test and the verification phase; (4) determine the role of the intensity of the verification phase (i.e., above or below the PO elicited at the end of the RI test) on the possible difference between the $\dot{V}O_2$ value during the RI test and the verification phase. We tested the hypothesis that a verification phase (either below or above the end-RI intensity) would produce a higher $\dot{V}O_2$ than that observed during the RI test.

MATERIALS AND METHODS

This study combines sets of data collected in two different locations: The University of Western Ontario and the University of Verona. Although part of the $\dot{V}O_{2max}$ data have been reported elsewhere (Murias et al., 2010b, 2011; Bruseghini et al., 2015; Keir et al., 2015) this is the first time that the $\dot{V}O_2$ data derived from the RI test and verification phase have been reported together and compared.

Participants

Data from a total of 61 healthy males (30 younger, Y: 25 ± 4 yr; 178 ± 6 cm; 79 ± 13 kg and 31 older adults O: 68 ± 5 yr; 174 ± 8 cm; 78 ± 10 kg; mean $\pm$ SD) were included in the present analysis. Eight younger and eight older adults performed the testing procedures at The University of Western Ontario (RI test + verification phase at 85% of peak PO at the end of the RI test). The remaining 22 younger and 23 older adults were tested at the University of Verona (RI test + verification phase at 105% of peak PO at the end of the RI test). All participants were volunteers and

provided written informed consent to participate in the study. Participants were included in this database if they were aged between 18 and 90 years old, had performed a RI test and a subsequent verification phase, were relatively naïve to laboratory testing procedures (i.e., had not been involved in laboratory testing in at least the previous 12 months). All participants were recreationally active, community dwelling individuals, none of whom was involved in an organized endurance training regime when the original testing took place. Common exclusion criteria were obesity (body mass index > 30 kg/m^2), smoking, involvement in any endurance training program within the previous twelve months, taking medications that would affect the cardiorespiratory or hemodynamic responses to exercise, or a history of cardiovascular, respiratory or musculoskeletal diseases. All subjects completed a PAR-Q+ questionnaire prior to enrolment and all older adults were medically screened by a physician before being accepted into the study. All procedures were approved by The University of Western Ontario Research Ethics Board for Health Sciences Research Involving Human Subjects or by the Ethics Board of the Department of Neurosciences, Biomedicine and Movement Sciences at the University of Verona.

Protocol

Participants performed a maximal RI test from a baseline of 20 W to the PO that elicited exhaustion using increments of 25 W/min (Y) and 15–20 W/min (O). Based on the anticipated fitness level of O (Pogliaghi et al., 2014) a ramp slope (i.e., W/min increment) was chosen that would bring the participants to exhaustion within 10–12 min based on the ACSM (2003) guidelines. The RI test was performed on a cycle ergometer (Lode Corival 400; Lode B.V., Groningen, Holland) and $\dot{V}O_2$ was measured throughout the test. After completion of the RI test, participants returned to cycling at the baseline PO of 20 W for 5 min, after which an instantaneous step increase in the PO was applied and the participants performed a constant load cycling exercise to volitional fatigue at a power output calculated to be either 85% ($n = 16$) or 105% ($n = 45$) of the peak PO achieved during the RI test. A similar approach has been previously described to assess a "true" plateau in the $\dot{V}O_2$ response and thus confirm the attainment of a true $\dot{V}O_{2max}$ (Sedgeman et al., 2013). By reducing the relative exercise intensity, the goal in this case was to obtain a longer time to exhaustion, in turn allowing sufficient time for a $\dot{V}O_{2max}$ to be achieved. The highest $\dot{V}O_2$ from each test was defined as the greatest consecutive 20-s average during each exhaustive test.

Measurements

Gas-exchange measurements at the University of Western Ontario were conducted as previously described (Babcock et al., 1994). Briefly, inspired and expired flow rates were measured using a low dead space (90 mL) bidirectional turbine (Alpha Technologies VMM 110) which was calibrated before tests using a syringe of known volume. Inspired and expired gases were sampled continuously (every 20 ms) at the mouth and analyzed for concentrations of O_2, CO_2, and nitrogen (N_2) by mass spectrometry (Perkin Elmer MGA-1100) after calibration

with precision-analyzed gas mixtures. Breath-by-breath alveolar gas exchange was calculated using the algorithms of Beaver et al. (1981). Heart rate (HR) was monitored continuously by a three-lead arrangement electrocardiogram using PowerLab (ML132/ML880; ADInstruments, Colorado Springs, CO). Data were recorded using LabChart v4.2 (ADInstruments, Colorado Springs, CO) on a separate computer.

At the University of Verona, breath-by-breath gas exchange and ventilation were continuously measured using a metabolic cart (QuarkB^2; COSMED, Rome, Italy) as previously described (Keir et al., 2015). The gas analyzers were calibrated before each experiment using a gas mixture of known concentration, and the turbine flowmeter was calibrated using a 3-L syringe (Hans Rudolph, Inc.). HR data were collected using radiotelemetry (SP0180 Polar Transmitter; Polar Electro, Inc., Kempele, Finland) and calculated over the duration of each breath.

Statistics

All statistics were performed using SPSS version 23 (SPSS, Chicago, IL). Data are presented as means $\pm$ SD. After ensuring the normal distribution of the data (Shapiro–Wilk test), and equality of variance for the 85% and 105% groups (Levene's test) a three-way repeated-measures ANOVA was used to test the possible influence and interactions of test type (RI test vs. constant load verification phase), age subgroup (Y vs. O) and verification phase intensity (85% vs. 105%) on the highest measures of V̇O$_2$ observed during the tests. Significant main effects and interactions were analyzed using the Bonferroni post-hoc test. Based on a power calculation, 49 participants were required to identify significant differences between groups for detecting changes larger than our estimated measurement error (i.e., 2 mL·kg^{-1}·min^{-1}) with a statistical power >0.8. The correlation between the highest V̇O$_2$ during the RI test and the constant load verification phase was assessed by Pearson's product moment correlation coefficient. Furthermore, the individual difference between V̇O$_2$ observed during the RI test the constant load verification phase was calculated in absolute units (absolute bias in mL·kg^{-1}·min^{-1}) and as a percent of RI V̇O$_2$ (% bias). Bland-Altman plots, followed by a one-sample z-test (Bland and Altman, 1986), were used to determine average bias, precision (i.e., standard deviation of the differences between measures) and limits of agreement between the V̇O$_2$ measures. The number and percent of subjects in which the RI test either overestimated or underestimated V̇O$_{2max}$ compared to the verification phase by over 2 mL·kg^{-1}·min^{-1} (equal to the minimum detectable change as measured in our laboratories for V̇O$_2$ measures between 2.1 and 3.5 L·min^{-1}) (Keir et al., 2014, 2015) was calculated. For all comparisons, statistical significance was declared when $P < 0.05$.

RESULTS

The peak PO was significantly lower in O (198 $\pm$ 35 W) compared to Y (339 $\pm$ 54 W; $P = 0.86$). Time to exhaustion during the RI test was shorter in O (10.2 $\pm$ 2.3 min; range 7.9–18.8 min) compared to Y (12.8 $\pm$ 2.0 min; range 8.9–18.6 min) ($p < 0.01$). Time to exhaustion during the verification phase was greater

when performing constant load exercise at 85% of peak PO (2.5 $\pm$ 0.4 min) compared to the constant load exercise performed at 105% of peak PO (1.7 $\pm$ 0.4 min) ($p < 0.01$). The average highest V̇O$_2$ values during the RI test in the whole group were not significantly different from the V̇O$_2$ values observed during the verification phase (**Table 1**; $P = 0.33$). This response was not affected by age, with no significant differences identified between the RI and verification V̇O$_2$ in O or Y (**Table 1**; $P = 0.64$). Similarly, there were no significant differences in the highest V̇O$_2$ values during RI test vs. verification phase at either 85% or 105% of peak PO ($P = 0.84$). The highest V̇O$_2$ values observed during both the RI test and the verification phase were highly correlated ($r = 0.99$; $p < 0.01$) (**Figure 1A**). The Bland-Altman analysis revealed a very small absolute bias (-0.25 mL·kg^{-1}·min^{-1}), which was not different from 0 ($z = -1.3$), and a precision of ±1.56 mL·kg^{-1}·min^{-1} between measures (**Figure 1B**). Bias $\pm$ precision was $-0.81 \pm 4.14\%$ when expressed as a percent of the RI values. Considering an absolute cut-off error of 2 mL·kg^{-1}·min^{-1}, the RI test overestimated V̇O$_{2max}$ compared to the verification phase in 4 participants (or ~7% of the group) and underestimated it in 6 participants (or ~10% of the group).

Peak HR observed during the RI test (171 $\pm$ 20 b·min^{-1}) was significantly higher than, and highly correlated with, the HR response observed during the verification phase (169 $\pm$ 20 b·min^{-1}; $p < 0.01$; $r = 0.96$; $p < 0.01$). Peak HR during both the RI test and the verification phase was significantly lower in O (153 $\pm$ 12 b·min^{-1} and 152 $\pm$ 14 b·min^{-1}, respectively) compared to Y (188 $\pm$ 8 b·min^{-1}; and 185 $\pm$ 9 b·min^{-1}, respectively; $p < 0.01$).

DISCUSSION

This study compared the V̇O$_2$ responses from a RI test to the limit of tolerance that was followed by a verification phase to exhaustion performed at intensities of either 85% or 105% of the peak PO observed at the end of the RI protocol, in a large and heterogeneous group of male participants that included older and younger individuals. The study tested whether the verification phase (either below or above the end-RI intensity) would produce a higher V̇O$_2$ than the maximal value observed during the RI test and whether such a difference would be affected by verification phase intensity and/or age. The main findings were that: (1) the average highest V̇O$_2$ values observed during the verification phases were not significantly higher than those obtained during the RI test; (2) the Bland-Altman analysis revealed a negligible bias (i.e., not different from zero); (3) the RI V̇O$_2$ was higher than the measurement error in V̇O$_2$ for ~7% and lower for ~10% of the group compared with the verification phase; (4) the difference between maximal V̇O$_2$ values associated with the verification phases compared to the RI tests was unaffected by the age of participants and verification phase intensity.

In the context of the unresolved uncertainties surrounding the achievement of a true V̇O$_{2max}$ during a RI test to exhaustion, Poole and Jones (2017) recently proposed that the term V̇O$_{2peak}$ should be avoided and that a verification phase, similar to that proposed by Rossiter et al. (2006) and further studied by others (Astorino et al., 2009; Barker et al., 2011; Sedgeman et al., 2013; Nolan et al., 2014; Sawyer et al., 2015), should

TABLE 1 | Highest $\dot{V}O_2$ and HR values during the ramp incremental (RI) test and the verification phase for each subgroup.

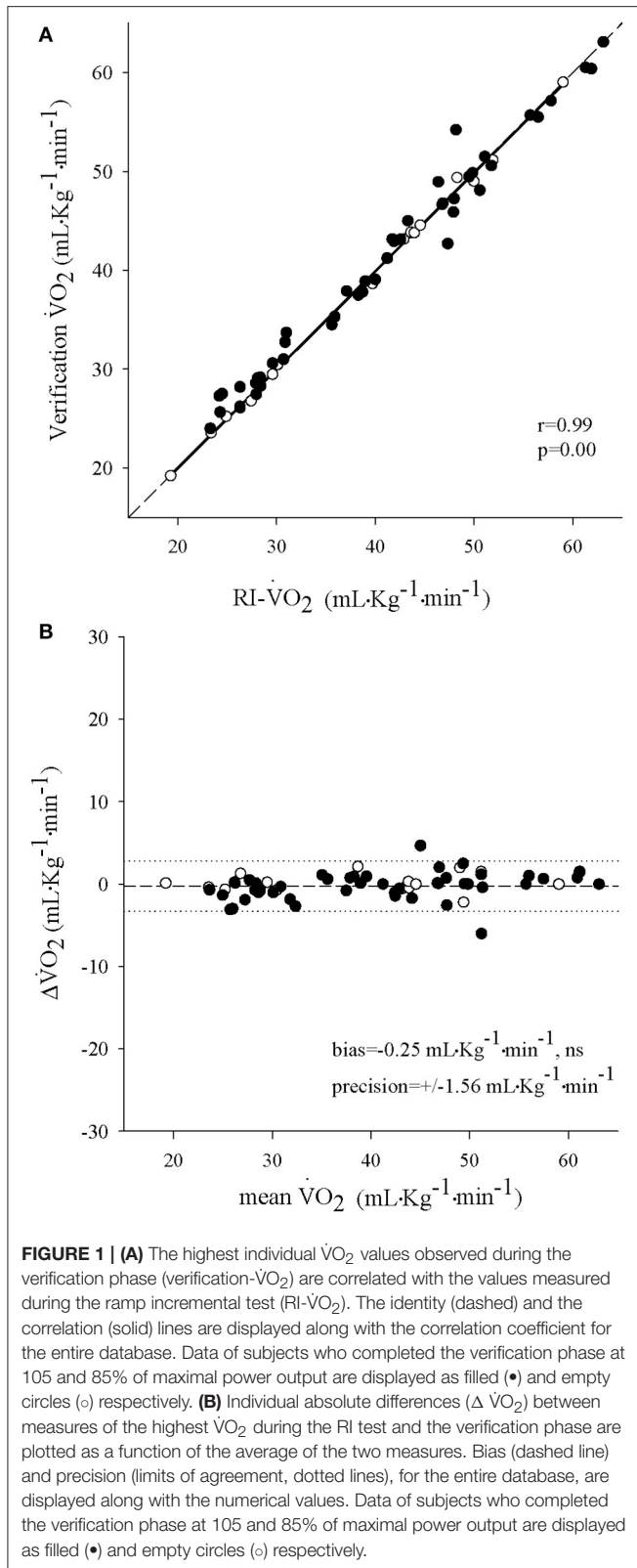| Entire data set	Total (n = 61)	Younger (n = 30)	Older (n = 31)
RI-$\dot{V}O_2$ (mL·kg^{-1}·min^{-1})	39.8 ± 11.5	48.8 ± 7.7	31.2 ± 7.1
Verification-$\dot{V}O_2$ (mL·kg^{-1}·min^{-1})	40.1 ± 11.2	48.8 ± 7.5	31.7 ± 6.9
RI-HR (b·min^{-1})	171 ± 20*	188 ± 8*	153 ± 12*
Verification-HR (b·min^{-1})	169 ± 20	185 ± 9	152 ± 14
Verification 85%	**Total (n = 16)**	**Younger (n = 8)**	**Older (n = 8)**
RI-$\dot{V}O_2$ (mL·kg^{-1}·min^{-1})	37.6 ± 11.9	47.2 ± 6.4	28.0 ± 7.1
Verification-$\dot{V}O_2$ (mL·kg^{-1}·min^{-1})	37.8 ± 11.9	47.6 ± 6.1	28.0 ± 7.0
RI-HR (b·min^{-1})	173 ± 24*	192 ± 6*	151 ± 14*
Verification-HR (b·min^{-1})	170 ± 24	189 ± 7	148 ± 15
Verification 105%	**Total (n = 45)**	**Younger (n = 22)**	**Older (n = 23)**
RI-$\dot{V}O_2$ (mL·kg^{-1}·min^{-1})	40.6 ± 11.4	49.4 ± 8.2	32.3 ± 6.9
Verification-$\dot{V}O_2$ (mL·kg^{-1}·min^{-1})	40.9 ± 10.9	49.2 ± 8.1	33.0 ± 6.5
RI-HR (b·min^{-1})	170 ± 19*	186 ± 8*	154 ± 11*
Verification-HR (b·min^{-1})	169 ± 19	184 ± 10	153 ± 13

*Average and standard deviation values of $\dot{V}O_2$ and HR as derived from the ramp incremental tests (RI) and from the verification trial in the whole group (Total) and in the young and older subgroups. Data from the entire dataset (that combines all tests irrespective of the validation trial intensity) are presented along with the separate averages of the two intensity subgroups. * Significantly different from the verification phase (p < 0.05).*

become a gold standard, especially in older, frail, and/or unfit populations. The idea that performing a verification phase at a given power output above that observed during the RI test to exhaustion is conceptually appealing as this procedure theoretically confirms that further increases in power output do not result in greater $\dot{V}O_2$ values (i.e., an upper limit plateau in the $\dot{V}O_2$ response). However, the evidence to support this type of statement appears surprisingly scant, and this idea neglects well-established concepts that show that exercising to the limit of tolerance at any intensity above critical power would result in attainment of $\dot{V}O_{2max}$.

Data from the present study indicated that, in 61 older and younger participants, the average highest $\dot{V}O_2$ values observed during the RI tests and the verification phase were similar, irrespective of age, fitness level and verification phase intensity. These findings confirm, in a larger and more heterogeneous population, the findings of others (Rossiter et al., 2006; Astorino et al., 2009; Barker et al., 2011; Sedgeman et al., 2013; Nolan et al., 2014; Sawyer et al., 2015). In addition, our study provides the first data of this type in an older adult population. Furthermore, individual differences between the RI test and the verification phase displayed quantitatively negligible (likely unmeasurable) and not significant bias. The results are similar to those in an athlete sample in whom the $\dot{V}O_2$ from RI and verification phase were not different, and within the error of measurement in 24 out of 24 subjects (Weatherwax et al., 2016).

According to the idea originally proposed by Rossiter et al. (2006) and examined by others (Astorino et al., 2009; Barker et al., 2011; Sedgeman et al., 2013; Nolan et al., 2014; Sawyer et al., 2015), and recently endorsed by Poole and Jones (2017), the lack of difference in the highest $\dot{V}O_2$ observed in RI vs. the verification phase would provide a proof of a plateau in $\dot{V}O_2$, in turn confirming the attainment of a true $\dot{V}O_{2max}$.

However, alternate and equally plausible views would be that: (1) the $\dot{V}O_{2max}$ obtained from the RI was "true" in the first place; (2) both the RI and the validation phase suffer from similar limitations (i.e., subjects' inability to endure maximal effort for a long enough time to allow oxidative metabolism to display a maximal functional level). In the first view, the verification phase would not add value beyond performing the RI test alone for providing a true measure of $\dot{V}O_{2max}$. In the second view, Poole and Jones (2017) highlighted that in less experienced participants or those perceived to be less willing to push themselves to their limit of tolerance, the RI model may result in a substantial underestimation of $\dot{V}O_{2max}$ and in such cases, a verification phase might offer an opportunity to highlight such underestimations. The present data set, as well as data from other studies in sedentary (Astorino et al., 2009), recreationally active (Sedgeman et al., 2013; Nolan et al., 2014), and overweight adults (Wood et al., 2010; Sawyer et al., 2015), as well as children (Barker et al., 2011) does not support the idea that a RI test underestimates $\dot{V}O_{2max}$ compared to a verification phase. These findings may in fact support a different interpretation: if $\dot{V}O_{2max}$ values derived from a RI test might underestimate the true $\dot{V}O_{2max}$ in less experienced/fit/healthy individuals, then these same people might not be able or willing to endure a maximal effort during a verification phase. Under these circumstances, the similar $\dot{V}O_2$ values observed in both tests might simply represent that participants are equally fatigued, experiencing tired legs or feeling somewhat breathless, but they still remain below their true $\dot{V}O_{2max}$ in both conditions. Thus, the verification phase presenting the same $\dot{V}O_2$ values as those observed during the RI test does not imply a plateau that necessarily reflects $\dot{V}O_{2max}$, or a verification of the RI test, but it might rather reflect the lack of its attainment in both tests. In this situation, true achievement of $\dot{V}O_{2max}$ is still uncertain but the "illusion" of

FIGURE 1 | (A) The highest individual V̇O$_2$ values observed during the verification phase (verification-V̇O$_2$) are correlated with the values measured during the ramp incremental test (RI-V̇O$_2$). The identity (dashed) and the correlation (solid) lines are displayed along with the correlation coefficient for the entire database. Data of subjects who completed the verification phase at 105 and 85% of maximal power output are displayed as filled (•) and empty circles (○) respectively. **(B)** Individual absolute differences (Δ V̇O$_2$) between measures of the highest V̇O$_2$ during the RI test and the verification phase are plotted as a function of the average of the two measures. Bias (dashed line) and precision (limits of agreement, dotted lines), for the entire database, are displayed along with the numerical values. Data of subjects who completed the verification phase at 105 and 85% of maximal power output are displayed as filled (•) and empty circles (○) respectively.

having measured a true V̇O$_{2max}$ is created. Given the difficulties associated with determination of a true V̇O$_{2max}$ value, and the limitations being highlighted with the verification phase as a

confirmatory test to establish V̇O$_{2max}$, the use of secondary markers of maximal effort become important. Although it has been suggested that secondary criteria might not be valid to provide accurate determination of the attainment of V̇O$_{2max}$ (Midgley et al., 2007; Poole et al., 2008), it could be argued that commonly used criteria such as [La] above 8 mmol·L^{-1}, a heart rate response within 10 bpm of the maximal predicted value, an RER higher than 1.10, and an RPE >18, might contribute to feel confidence in the attainment of a true V̇O$_{2max}$ at the end of a RI test.

Perhaps the recommendation for the utilization of a verification phase should recognize that the aim of this approach is not to identify V̇O$_{2max}$ in a majority of participants, but to find the minority of participants who might not achieve it during a RI test. Nevertheless, with a precision of ~1.5 mL·kg^{-1}·min^{-1} within the minimum detectable change of 2 mL·kg^{-1}·min^{-1}, indeed the verification phase revealed very few instances of an underestimation from the RI. Interestingly, the data from the present study showed that the V̇O$_2$ values during the RI test similarly over- (i.e., four tests) and underestimated (i.e., six tests) the V̇O$_2$ values associated with the verification phase, suggesting major limitations in this verification procedure.

In relation to the effectiveness of the verification phase to reach V̇O$_{2max}$, the relationship between the intensity and the duration of the exercise should be considered. In the present study, the times to exhaustion for the verification phases performed five min after the end of the RI test were ~2.5 and ~1.5 min for exercise performed at 85% and 105% of the peak PO, respectively. Although this time might be considered too short for V̇O$_{2max}$ to develop (especially for more intense and shorter exercise bouts) (Poole and Jones, 2012), the fact that the verification phases were performed shortly after the end of the RI test and that the system was "primed" (De Roia et al., 2012) might have contributed to the achievement of a V̇O$_2$ response that was as high as that seen toward the end of the RI test. While some papers have used supramaximal intensities of exercise for the verification phases in order to establish a "plateau response" despite an increase in PO beyond that observed during the RI test (Astorino et al., 2009; Barker et al., 2011; Nolan et al., 2014), Sedgeman et al. (2013) compared a verification phase performed at 105% of the peak PO during the RI test to a verification phase performed at a PO equivalent to that occurring two stages before the end of a graded test (i.e., ~54 W lower PO or ~82% of peak PO), subsequent to 3 min recovery after the end of the RI test. As found for the present data using 85% or 105% of the end-RI work rate, Sedgeman et al. (2013) also showed the V̇O$_2$ values were not different between the RI test and the verification phase either with the sub-maximal (~2.2 min duration) or the supramaximal (~1.3 min duration) verification. These authors suggested that a sub-maximal verification phase might be beneficial as more plateau responses were observed during this type of verification. This interpretation is not surprising; exercising within the very heavy intensity domain (such as in Sedgeman et al. and the present studies) is compatible with a longer exercise duration, in turn allowing enough time for the slow component of V̇O$_2$ to develop and for V̇O$_{2max}$ to be achieved. On the contrary, exercising in the severe intensity domain (e.g., 105% of peak

PO) might pose some extra challenges for confirmation of a true $\dot{V}O_{2max}$ as a plateau response is less likely to occur and, although not the case in the present study and in that of Sedgeman et al., could theoretically result in muscle fatigue before the $\dot{V}O_2$ response can be fully developed and $\dot{V}O_{2max}$ expressed (Poole and Jones, 2012). This is an important aspect to consider as the duration of recovery period before the verification phase and/or the characteristics of the ramp during the incremental test to exhaustion (i.e., steeper ramps will result in higher peak POs compared to less steep ramps) might determine whether or not the highest possible $\dot{V}O_2$ is achieved when the intensity of the verification phase is supramaximal.

In the present study, the verification phase at 85% and 105% of peak PO were collected in two different laboratories, which might represent a possible limitation. However, the equipment used in both laboratories was comparably accurate and precise and the experimental procedures were consistent between laboratories and clearly described to allow replication. Furthermore, since the analysis compares $\dot{V}O_2$ values obtained with the RI test and verification phase within each individual, whatever systematic bias in $\dot{V}O_2$ measures might exist between laboratories would affect both data-points in the same way. As for the vast majority of the studies that have described the use of a verification phase, the population tested in the present study is limited to male participants. It was important to verify the usefulness of the verification phase in the same population in which the verification phase has been proposed; however, the absence of female participants in this study is an issue that, although not expected to change the outcome observed in the present investigation, might need to be addressed in future studies.

In conclusion, the accurate identification of $\dot{V}O_{2max}$ remains an elusive issue and there is a clear need for developing reliable criteria to objectively evaluate the achievement of a true maximal response. In this context, the proposal by Poole and Jones (2017) to use the verification phase for true $\dot{V}O_{2max}$ determination appeared a very promising approach to overcome the absence of an objective plateau in $\dot{V}O_2$. However, data presented in this study indicate that a verification phase performed subsequent to the end of a RI test does not provide a convincing confirmation that a true $\dot{V}O_{2max}$ response has been achieved. Therefore, the recommendation that a verification phase should become the gold standard procedure for $\dot{V}O_{2max}$ determination, although initially appealing, is not supported. As such, the verification phase should not be accepted as a gold standard, and given the present analysis indicating that in most cases the $\dot{V}O_2$ during the RI test is as high as that of the verification phase, or that the highest $\dot{V}O_2$ in the few cases observed on either the RI test or the verification phase is within the error of the measurement, further efforts to endorse the verification phase to confirm $\dot{V}O_{2max}$ do not seem to be a tenable direction for future research.

AUTHOR CONTRIBUTIONS

JM, SP, and DP: contributed to the conception of the work, analysis, and interpretation of data; JM, SP, and DP: contributed to the drafting and revisions of the manuscript; JM, SP, and DP: approved the final version of this manuscript; JM, SP, and DP: agree to be accountable for all aspects of the work in ensuring that questions related to the accuracy or integrity of any part of the work are appropriately investigated and resolved.

FUNDING

ACKNOWLEDGMENTS

REFERENCES

ACSM (2003). *ACSM's Guidelines for Exercise Testing and Prescription.* Philadelphia, PA: Lippincot Williams and Wilkins.

Astorino, T. A., White, A. C., and Dalleck, L. C. (2009). Supramaximal testing to confirm attainment of $\dot{V}O_{2max}$ in sedentary men and women. *Int. J. Sports Med.* 30, 279–284. doi: 10.1055/s-0028-1104588

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

Babcock, M. A., Paterson, D. H., and Cunningham, D. A. (1994). Effects of aerobic endurance training on gas exchange kinetics of older men. *Med. Sci. Sports Exerc.* 26, 447–452.

Barker, A. R., Williams, C. A., Jones, A. M., and Armstrong, N. (2011). Establishing maximal oxygen uptake in young people during a ramp cycle test to exhaustion. *Br. J. Sports Med.* 45, 498–503. doi: 10.1136/bjsm.2009.063180

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

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

Borrelli, E., Pogliaghi, S., Molinello, A., Diciolla, F., Maccherini, M., and Grassi, B. (2003). Serial assessment of peak $\dot{V}O_2$ and $\dot{V}O_2$ kinetics early after heart transplantation. *Med. Sci. Sports Exerc.* 35, 1798–1804. doi: 10.1249/01.MSS.0000093610.71730.02

Bowen, T. S., Cannon, D. T., Begg, G., Baliga, V., Witte, K. K., and Rossiter, H. B. (2012). A novel cardiopulmonary exercise test protocol and criterion to determine maximal oxygen uptake in chronic heart failure. *J. Appl. Physiol.* 113, 451–458. doi: 10.1152/japplphysiol.01416.2011

Bringard, A., Pogliaghi, S., Adami, A., De Roia, G., Lador, F., Lucini, D., et al. (2010). Cardiovascular determinants of maximal oxygen consumption in upright and supine posture at the end of prolonged bed rest in humans. *Respir. Physiol. Neurobiol.* 172, 53–62. doi: 10.1016/j.resp.2010.03.018

Bruseghini, P., Calabria, E., Tam, E., Milanese, C., Oliboni, E., Pezzato, A., et al. (2015). Effects of eight weeks of aerobic interval training and of isoinertial resistance training on risk factors of cardiometabolic diseases in healthy elderly subjects. *Oncotarget* 6, 16998–17015. doi: 10.18632/oncotarget.4031

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. Respir. Environ. Exerc. Physiol.* 55, 1558–1564.

Day, J. R., Rossiter, H. B., Coats, E. M., Skasick, A., and Whipp, B. J. (2003). The maximally attainable $\dot{V}O_2$ during exercise in humans: the peak vs. maximum issue. *J. Appl. Physiol.* 95, 1901–1907. doi: 10.1152/japplphysiol.00024.2003

De Roia, G., Pogliaghi, S., Adami, A., Papadopoulou, C., and Capelli, C. (2012). Effects of priming exercise on the speed of adjustment of muscle oxidative metabolism at the onset of moderate-intensity step transitions in older adults. *Am. J. Physiol. Regul. Integr. Comp. Physiol.* 302, R1158–R1166. doi: 10.1152/ajpregu.00269.2011

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

Doria, C., Toniolo, L., Verratti, V., Cancellara, P., Pietrangelo, T., Marconi, V., et al. (2011). Improved $\dot{V}O_2$ uptake kinetics and shift in muscle fiber type in high-altitude trekkers. *J. Appl. Physiol.* 111, 1597–1605. doi: 10.1152/japplphysiol.01439.2010

Frontera, W. R., Meredith, C. N., O'Reilly, K. P., and Evans, W. J. (1990). Strength training and determinants of $\dot{V}O_{2max}$ in older men. *J. Appl. Physiol.* 68, 329–333.

Glassford, R. G., Baycroft, G. H., Sedgwick, A. W., and Macnab, R. B. (1965). Comparison of maximal oxygen uptake values determined by predicted and actual methods. *J. Appl. Physiol.* 20, 509–513.

Hill, A. V., and Lupton, H. (1923). Muscular exercise, lactic acid and the supply and utilization of oxygen. *Q. J. Med.* 16, 135–171.

Hoppeler, H., and Weibel, E. R. (2000). Structural and functional limits for oxygen supply to muscle. *Acta Physiol. Scand.* 168, 445–456. doi: 10.1046/j.1365-201x.2000.00696.x

Jensen, M. T., Holtermann, A., Bay, H., and Gyntelberg, F. (2016). Cardiorespiratory fitness and death from cancer: a 42-year follow-up from the Copenhagen male study. *Br. J. Sports Med.* 51, 1364–1369. doi: 10.1136/bjsports-2016-096860

Keir, D. A., Fontana, F. Y., Robertson, T. C., Murias, J. M., Paterson, D. H., Kowalchuk, J. M., et al. (2015). Exercise intensity thresholds: identifying the boundaries of sustainable performance. *Med. Sci. Sports Exerc.* 47, 1932–1940. doi: 10.1249/MSS.0000000000000613

Keir, D. A., Murias, J. M., Paterson, D. H., and Kowalchuk, J. M. (2014). Breath-by-breath pulmonary O_2 uptake kinetics: effect of data processing on confidence in estimating model parameters. *Exp. Physiol.* 99, 1511–1522. doi: 10.1113/expphysiol.2014.080812

Krogh, A., and Lindhard, J. (1913). The regulation of respiration and circulation during the initial stages of muscular work. *J. Physiol.* 47, 112–136.

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

McGawley, K. (2017). The reliability and validity of a four-minute running time-trial in assessing $\dot{V}O_{2max}$ and performance. *Front. Physiol.* 8:270. doi: 10.3389/fphys.2017.00270

Midgley, A. W., and Carroll, S. (2009). Emergence of the verification phase procedure for confirming 'true' $\dot{V}O_{2max}$. *Scand. J. Med. Sci. Sports* 19, 313–322. doi: 10.1111/j.1600-0838.2009.00898.x

Midgley, A. W., McNaughton, L. R., Polman, R., and Marchant, D. (2007). Criteria for determination of maximal oxygen uptake: a brief critique and recommendations for future research. *Sports Med.* 37, 1019–1028. doi: 10.2165/00007256-200737120-00002

Murias, J. M., Kowalchuk, J. M., and Paterson, D. H. (2010a). Mechanisms for increases in $\dot{V}O_{2max}$ with endurance training in older and young women. *Med. Sci. Sports Exerc.* 42, 1891–1898. doi: 10.1249/MSS.0b013e3181dd0bba

Murias, J. M., Kowalchuk, J. M., and Paterson, D. H. (2010b). Time course and mechanisms of adaptations in cardiorespiratory fitness with endurance training in older and young men. *J. Appl. Physiol.* 108, 621–627. doi: 10.1152/japplphysiol.01152.2009

Murias, J. M., Kowalchuk, J. M., Ritchie, D., Hepple, R. T., Doherty, T. J., and Paterson, D. H. (2011). Adaptations in capillarization and citrate synthase activity in response to endurance training in older and young men. *J. Gerontol. A Biol. Sci. Med. Sci.* 66, 957–964. doi: 10.1093/gerona/glr096

Nolan, P. B., Beaven, M. L., and Dalleck, L. (2014). Comparison of intensities and rest periods for $\dot{V}O_{2max}$ verification testing procedures. *Int. J. Sports Med.* 35, 1024–1029. doi: 10.1055/s-0034-1367065

Paterson, D. H., and Warburton, D. E. (2010). Physical activity and functional limitations in older adults: a systematic review related to Canada's physical activity guidelines. *Int. J. Behav. Nutr. Phys. Act.* 7:38. doi: 10.1186/1479-5868-7-38

Pogliaghi, S., Bellotti, C., and Paterson, D. H. (2014). "Tailored" submaximal step test for $\dot{V}O_{2max}$ prediction in healthy older adults. *J. Aging Phys. Act.* 22, 261–268. doi: 10.1123/japa.2012-0171

Pogliaghi, S., Terz'iotti, P., Cevese, A., Balestreri, F., and Schena, F. (2006). Adaptations to endurance training in the healthy elderly: arm cranking versus leg cycling. *Eur. J. Appl. Physiol.* 97, 723–731. doi: 10.1007/s00421-006-0229-2

Poole, D. C., and Jones, A. M. (2012). Oxygen uptake kinetics. *Compr. Physiol.* 2, 933–996. doi: 10.1002/cphy.c100072

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

Poole, D. C., Wilkerson, D. P., and Jones, A. M. (2008). Validity of criteria for establishing maximal O_2 uptake during ramp exercise tests. *Eur. J. Appl. Physiol.* 102, 403–410. doi: 10.1007/s00421-007-0596-3

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

Sawyer, B. J., Tucker, W. J., Bhammar, D. M., and Gaesser, G. A. (2015). Using a verification test for determination of $\dot{V}O_{2max}$ in sedentary adults with obesity. *J. Strength Cond. Res.* 29, 3432–3438. doi: 10.1519/JSC.0000000000001199

Sedgeman, D., Dalleck, L., Clark, I. E., Jamnick, N., and Pettitt, R. W. (2013). Analysis of square-wave bouts to verify $\dot{V}O_{2max}$. *Int. J. Sports Med.* 34, 1058–1062. doi: 10.1055/s-0033-1341436

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.

Weatherwax, R. M., Richardson, T. B., Beltz, N. M., Nolan, P. B., and Dalleck, L. (2016). Verification testing to confirm $\dot{V}O_{2max}$ in altitude-residing, endurance-trained runners. *Int. J. Sports Med.* 37, 525–530. doi: 10.1055/s-0035-1569346.

Wood, R. E., Hills, A. P., Hunter, G. R., King, N. A., and Byrne, N. M. (2010). $\dot{V}O_{2max}$ in overweight and obese adults: do they meet the threshold criteria? *Med. Sci. Sports Exerc.* 42, 470–477. doi: 10.1249/MSS.0b013e3181b666ad