Medicine & Science IN Sports & Exercise

The Official Journal of the American College of Sports Medicine
www.acsm-msse.org

. . . Published ahead of Print

Maximal Oxygen Uptake Is Underestimated during Incremental Testing in Hypertensive Older Adults: Findings from the HAEL Study

Gustavo Z. Schaun[1,2], Cristine L. Alberton[1], Maria Laura B. Gomes[1], Lucas P. Santos[3,4], Marcas M. Bamman[2,5]; Graciele F. Mendes[1], Mariana S. Häfele[1], Luana S. Andrade[1], Leonardo Alves[6], Vinícius A. de Ataides[7]; Marco A. Carmona[7]; Rafael Lázaro[7]; Cíntia E. Botton[1,3,4], Daniel Umpierre[3,4,8], Stephanie S. Pinto[1], Eurico N. Wilhelm[1,9]

[1]Neuromuscular Assessment Laboratory, Physical Education School, Federal University of Pelotas, Pelotas, RS, Brazil; [2]UAB Center for Exercise Medicine, University of Alabama at Birmingham, Birmingham, AL; [3]Exercise Pathophysiology Laboratory, Centro de Pesquisa Clínica, Hospital de Clínicas de Porto Alegre, Universidade Federal do Rio Grande do Sul, Porto Alegre, RS, Brazil; [4]National Institute of Science and Technology for Health Technology Assessment (IATS/HCPA), Hospital de Clínicas de Porto Alegre, Porto Alegre, RS, Brazil; [5]Departments of Cell, Developmental, and Integrative Biology; Medicine; and Neurology, University of Alabama at Birmingham, Birmingham, AL; [6]Medical School, Federal University of Rio Grande, Rio Grande, RS, Brazil; [7]Departamento de Nefrologia, Hospital Universitário São Francisco de Paula, Catholic University of Pelotas, Pelotas, RS, Brazil; [8]Department of Public Health, Universidade Federal do Rio Grande do Sul, Porto Alegre, RS, Brazil; [9]Department of Sport, Exercise and Rehabilitation, Northumbria University, Newcastle upon Tyne, Tyne and Wear, United Kingdom

Accepted for Publication: 15 December 2020

ABSTRACT

Purpose: The present cross-sectional study aimed to investigate whether a maximal oxygen uptake (VO_{2max}) verification phase (VER) could improve the accuracy of a previous graded exercise test (GXT) to assess individual VO_{2max} in hypertensive individuals. **Methods:** Thirty-three older adults with hypertension (24 women) taking part in the Hypertension Approaches in the Elderly Study (NCT03264443) were recruited. Briefly, after performing a treadmill GXT to exhaustion, participants rested for 10 min and underwent a multistage VER to confirm GXT results. Individual VO_{2max}, respiratory exchange ratio (RER), maximal heart rate (HR_{max}), and rating of perceived exertion (RPE) were measured during both GXT and VER tests. Mean values were compared between bouts using paired sample t-tests and VO_{2max} was also compared between GXT and VER on an individual basis. **Results:** Testing was well tolerated by all participants. Both absolute ($p=0.011$) and relative ($p=0.014$) VO_{2max} values were higher in VER than in GXT. RER ($p<0.001$) and RPE ($p=0.002$) were lower in VER, whereas HR_{max} ($p=0.286$) was not different between the two trials. Individual VO_{2max} comparisons revealed that 54.6% of the participants (18/33) achieved a VO_{2max} value that was ≥3% during VER (mean: 13.5%, range: from +3% to +22.1%, ES=0.062), whereas 87.9% (29/33) of the tests would have been validated as a maximal effort if the classic criteria were employed (i.e. VO_2 plateau or at least two secondary criteria). **Conclusion:** In sedentary older individuals with hypertension, GXT to exhaustion underestimated VO_{2max} in more than half of tested participants, even when established but criticized criteria were used to confirm whether a maximal effort was attained. Employing VER after GXT is a quick approach to assist with the verification of an individual's VO_{2max}.

Maximal Oxygen Uptake Is Underestimated during Incremental Testing in Hypertensive Older Adults: Findings from the HAEL Study

Gustavo Z. Schaun[1,2], Cristine L. Alberton[1], Maria Laura B. Gomes[1], Lucas P. Santos[3,4], Marcas M. Bamman[2,5]; Graciele F. Mendes[1], Mariana S. Häfele[1], Luana S. Andrade[1], Leonardo Alves[6], Vinícius A. de Ataides[7]; Marco A. Carmona[7]; Rafael Lázaro[7]; Cíntia E. Botton[1,3,4], Daniel Umpierre[3,4,8], Stephanie S. Pinto[1], Eurico N. Wilhelm[1,9]

[1]Neuromuscular Assessment Laboratory, Physical Education School, Federal University of Pelotas, Pelotas, RS, Brazil; [2]UAB Center for Exercise Medicine, University of Alabama at Birmingham, Birmingham, AL; [3]Exercise Pathophysiology Laboratory, Centro de Pesquisa Clínica, Hospital de Clínicas de Porto Alegre, Universidade Federal do Rio Grande do Sul, Porto Alegre, RS, Brazil; [4]National Institute of Science and Technology for Health Technology Assessment (IATS/HCPA), Hospital de Clínicas de Porto Alegre, Porto Alegre, RS, Brazil; [5]Departments of Cell, Developmental, and Integrative Biology; Medicine; and Neurology, University of Alabama at Birmingham, Birmingham, AL; [6]Medical School, Federal University of Rio Grande, Rio Grande, RS, Brazil; [7]Departamento de Nefrologia, Hospital Universitário São Francisco de Paula, Catholic University of Pelotas, Pelotas, RS, Brazil; [8]Department of Public Health, Universidade Federal do Rio Grande do Sul, Porto Alegre, RS, Brazil; [9]Department of Sport, Exercise and Rehabilitation, Northumbria University, Newcastle upon Tyne, Tyne and Wear, United Kingdom

Corresponding author
Gustavo Z. Schaun
E-mail: gustavoschaun@hotmail.com
Tel: +55 53 3273-2752
Fax: +55 53 3273-3851
Address: Rua Luís de Camões, 625; Tablada, Pelotas-RS, Brazil. 96055-630.

Running head: VO_{2max} verification in hypertensive patients

The HAEL Study or involved personnel have been funded by CNPq, grant number 429849/2016-8 (Conselho Nacional de Desenvolvimento Científico e Tecnológico), FAPERGS (Fundação de Amparo à Pesquisa do Estado do Rio Grande do Sul), CAPES (Coordenação de Aperfeiçoamento de Pessoal de Nível Superior), and FIPE (Fundo de Incentivo à Pesquisa e Eventos, Hospital de Clinicas de Porto Alegre). The authors are thankful for the administrative and funding support provided by the National Institute of Science and Technology for Health Technology Assessment (IATS/HCPA), Hospital de Clínicas de Porto Alegre, Porto Alegre, RS, Brazil. E. N. W. was supported by the Brazilian Education Ministry Foundation CAPES (PNPD). **Conflicts of Interest.** All authors declare that they have no competing financial interests related to this publication. All authors declare that the results of the study are presented clearly, honestly, and without fabrication, falsification, or inappropriate data manipulation. The results of the present study do not constitute an endorsement by the American College of Sports Medicine.

ABSTRACT

Purpose: The present cross-sectional study aimed to investigate whether a maximal oxygen uptake (VO_{2max}) verification phase (VER) could improve the accuracy of a previous graded exercise test (GXT) to assess individual VO_{2max} in hypertensive individuals. **Methods:** Thirty-three older adults with hypertension (24 women) taking part in the Hypertension Approaches in the Elderly Study (NCT03264443) were recruited. Briefly, after performing a treadmill GXT to exhaustion, participants rested for 10 min and underwent a multistage VER to confirm GXT results. Individual VO_{2max}, respiratory exchange ratio (RER), maximal heart rate (HR_{max}), and rating of perceived exertion (RPE) were measured during both GXT and VER tests. Mean values were compared between bouts using paired sample t-tests and VO_{2max} was also compared between GXT and VER on an individual basis. **Results:** Testing was well tolerated by all participants. Both absolute ($p=0.011$) and relative ($p=0.014$) VO_{2max} values were higher in VER than in GXT. RER ($p<0.001$) and RPE ($p=0.002$) were lower in VER, whereas HR_{max} ($p=0.286$) was not different between the two trials. Individual VO_{2max} comparisons revealed that 54.6% of the participants (18/33) achieved a VO_{2max} value that was ≥3% during VER (mean: 13.5%, range: from +3% to +22.1%, ES=0.062), whereas 87.9% (29/33) of the tests would have been validated as a maximal effort if the classic criteria were employed (i.e. VO_2 plateau or at least two secondary criteria). **Conclusion:** In sedentary older individuals with hypertension, GXT to exhaustion underestimated VO_{2max} in more than half of tested participants, even when established but criticized criteria were used to confirm whether a maximal effort was attained. Employing VER after GXT is a quick approach to assist with the verification of an individual's VO_{2max}.

Keywords: Exercise testing; Verification phase; Oxygen consumption; Hypertension; Blood pressure

INTRODUCTION

Maximal oxygen uptake ($\dot{V}O_{2max}$) represents the maximum rate at which the body can transport and utilize oxygen for cellular work (1). This concept was initially introduced by Hill and Lupton (2), and since then $\dot{V}O_{2max}$ has become one of the most comprehensive measures of cardiorespiratory capacity (3), being considered an important health marker (4) and an independent predictor of all-cause mortality and cardiovascular events (4, 5). The assessment of $\dot{V}O_{2max}$, however, has been a matter of debate and several criteria have been proposed to confirm whether the oxygen uptake ($\dot{V}O_2$) obtained at the end of a maximal incremental effort is representative of the individual's maximum capacity to consume oxygen. Among them, the most known and possibly most used criterion is the presence of a $\dot{V}O_2$ plateau (6). Although considered the gold standard for $\dot{V}O_{2max}$ determination, a plateau in $\dot{V}O_2$ is not always observed, ranging from 0% to 100%, and is influenced by other factors such as the exercise mode and test protocol employed (3, 7), as well as data analysis and interpretation methods (1). Consequently, reliance solely on the plateau criterion has been criticized (3, 8–11).

In the absence of a plateau, secondary criteria have been used as an indication that a maximal effort was attained during graded exercise testing (GXT) (12). These include attaining age-estimated maximal heart rate (HR), respiratory exchange ratio (RER) > 1.1 or 1.15, and Borg's rating of perceived exertion (RPE) values (8, 9). Although previous researchers have stated that the presence of at least two of these criteria is a strong indicator that $\dot{V}O_{2max}$ was attained (8, 12), others have criticized its capacity to discriminate a "true" $\dot{V}O_{2max}$ (1, 10, 13), with some authors even suggesting a complete rejection of secondary criteria (14, 15). As a consequence, alternative strategies have been proposed, such as the so-called verification phase (VER) (9, 10), which is

characterized as an additional effort performed above critical power, or ideally above the work rate that yielded $\dot{V}O_{2max}$, after GXT itself (16). Lack of a difference between the $\dot{V}O_2$ at the end of GXT and VER is perhaps an ideal method for determining true $\dot{V}O_{2max}$ was reached during GXT (14–17).

In clinical settings, GXT has been typically confirmed using traditional criteria (i.e. $\dot{V}O_2$ plateau and secondary criteria) in a symptom-limited manner. However, this practice has also been recently criticized (17) and, based on the low incidence of plateaus and the possibility that secondary criteria underestimate $\dot{V}O_{2max}$ it makes the search for alternatives necessary and relevant. Within this context, VER has been previously tested and found valid for use in physically active elderly men and women (18), but its applicability and benefit for clinical populations deserve further investigation.

Systemic arterial hypertension is a worldwide highly prevalent condition, with global estimates suggesting about 31% of the adult population is hypertensive (19). High cardiorespiratory fitness is an important factor associated with a lower risk of death from cardiovascular diseases, including hypertension (5). As such, adequate $\dot{V}O_{2max}$ assessment is essential to both clinical and laboratory practice to support proper diagnosis and prediction of cardiovascular risk, develop individual recommendations about physical activity and exercise, as well as to track and follow-up training-related outcomes (20–22). To date, however, it is unknown whether VER is both feasible and provides additional information to confirm $\dot{V}O_{2max}$ in clinical populations, such as hypertensive individuals. Considering the prevalence of hypertension in older adults, which may exceed 70% in those aged 60 and over (23), hypertensive patients were used in the present investigation as a

model for individuals and/or conditions that are typically associated to lower exercise tolerance or fitness levels as compared to healthy or physically active older adults, that presents. Therefore, the aim of this study was to investigate whether VER was able to confirm the $\dot{V}O_{2max}$ results of a GXT in older adults with hypertension. Furthermore, we compared VER with other criteria previously employed to validate a maximum GXT (e.g. $\dot{V}O_2$ plateau, cutoff point in RER, estimated maximum HR and near-maximum RPE). Our hypothesis was that VER would improve the determination of $\dot{V}O_{2max}$ in the studied population compared to other criteria.

METHODS

This investigation is a cross-sectional parallel study based on the *Hypertension Approaches in the Elderly: a Lifestyle Study* (HAEL) trial, a multicenter randomized clinical trial that compared the effects of pragmatic combined exercise training vs. a health education program on ambulatory blood pressure and cardiorespiratory fitness in older adults with hypertension (ClinicalTrials.gov NCT03264443). The current manuscript focuses on the Pelotas center participants' maximal GXT and VER tests and $\dot{V}O_2$ data at baseline, which were carried out from February 2018 to June 2019 and includes all but the first recruitment wave of the study. For further details of the HAEL Study please refer to the protocol manuscript (24).

Participants

Out of the forty-two individuals included in the HAEL Study during the recruitment period of the current investigation, two did not perform GXT, six refused to participate in the VER study, and data from one participant could not be retrieved due to equipment malfunctioning during VER. Thus, thirty-three participants (9 men and 24 women) ≥60 years old previously diagnosed with

hypertension agreed to participate and were included in the present analysis (Table 1). Briefly, exclusion criteria followed those of the HAEL Study and included a history of myocardial infarction, myocardial revascularization, deep vein thrombosis, stroke or pulmonary embolism in the last 12 months; severe cardiovascular disease (e.g. class III or IV chronic heart failure or unstable arrhythmia); chronic lung disease requiring O_2 or corticosteroids; kidney disease requiring dialysis; neurological disorders; cancer requiring treatment in the last two years or medical contraindication to exercising. Prior to any procedure, all participants read and signed an informed consent form and the research project was previously approved by the Federal University of Pelotas Institutional Review Board (CAAE: 62427616.0.2001.5313) in accordance with the Declaration of Helsinki.

Insert Table 1 near here

Experimental procedures

The present study was conducted during the baseline maximal GXT session of the HAEL Study (24). Specifically, participants performed a GXT to determine their $\dot{V}O_{2max}$ followed by VER (explained in detail below) after a brief recovery period. For safety reasons, all tests were supervised by a physician and participants were familiarized with the equipment and safety procedures prior to the experimental session.

Protocols

Maximal incremental test. Participants attended the laboratory and, after a 5-min resting interval, baseline resting blood pressure measurement was taken and the participants were presented to the

physician along with a brief case summary of the patient. After clearance, participants were submitted to a maximal GXT to exhaustion on a motorized treadmill. Participants warmed-up at 3 $km \cdot h^{-1}$ for 3 min, and the test started with 0.5 $km \cdot h^{-1}$ and 1% grade increments per minute until volitional failure or at the physician's request. For safety, after every two minutes, and at least 30 seconds after the workload was increased, measurements of blood pressure and RPE (Borg 6-20) were obtained, and the participant's HR was recorded. All participants were instructed to have their last meal at least 3 hours before the test, and to abstain from intense physical exercises and ingestion of caffeine and alcohol 24 hours before the experimental session.

Verification Phase. After a 10-min passive recovery interval, each participant underwent a VER to determine whether the $\dot{V}O_{2max}$ value obtained during GXT was a true maximum. Specifically, a new multistage effort bout was performed, which consisted of 2 min at 50% of the maximal speed and grade reached during GXT, followed by 1 min at 70% of the maximal speed and grade at GXT, and then participants exercised to volitional failure at one stage above the last completed stage in their individual GXT (16, 25). We chose a multistage VER rather than a single square-wave protocol to minimize the chance of patients reaching fatigue before $\dot{V}O_{2max}$ was attained due to the possibility of slow $\dot{V}O_2$ kinetics (26).

During both GXT and VER, gas exchange data were obtained using a portable metabolic system (VO2000, MedGraphics®, Ann Arbor, MI, USA), which was calibrated prior to each session according to manufacturer's specifications and recalibrated between GXT and VER to minimize drift. Participant's HR was measured telemetrically using a HR monitor (Polar FT1, Polar Electro, Finland), and the exercise test would be terminated if adverse symptoms were observed, such as

blood pressure above 250/115 mm Hg, a decline in systolic blood pressure with increasing work rate or at the physician's discretion (20), although no test had to be interrupted because of these reasons.

Data analysis

$\dot{V}O_2$ values were time-averaged every 20 s and the average from the last two 20-s intervals in both GXT and VER were considered as phase-specific $\dot{V}O_{2max}$ (13, 27). The $\dot{V}O_{2max}$ during GXT was confirmed when the difference between VER and GXT $\dot{V}O_{2max}$ was less than 3% (25, 28). In addition, to determine the number of valid tests, GXT results were also confirmed based on the presence of a $\dot{V}O_2$ plateau ($\dot{V}O_2$ change $\leq$ 150 $ml \cdot min^{-1}$ with an increase in workload); or when at least two secondary criteria were met (i.e. ± 10 bpm of age-predicted HR ($\geq$ 220-age), RER > 1.10 or RPE $\geq$ 18) (6, 20). To gain insight into GXT capacity to confirm $\dot{V}O_{2max}$ in relation to VER, a novel individualized plateau criterion proposed by Midgley et al. (25) was also used in which a least squares linear regression line was fitted to the linear portion of GXT $\dot{V}O_2$-workrate data excluding the final 2 min. The regression line was then extrapolated to the end of the test and a plateau was confirmed when a difference between GXT modeled and measured $\dot{V}O_{2max}$ was greater than 50% of the regression slope for the linear portion (25).

Statistical Analysis

Data are reported as mean ± standard deviation (SD) or confidence intervals (CI) unless stated otherwise. Initially, comparisons between the GXT and VER mean outcomes were performed using paired sample t-tests. Considering that the comparison between mean values may conceal between-trials variability, comparisons between GXT and VER $\dot{V}O_{2max}$ results were also

conducted individually based on the $\dot{V}O_{2max}$ difference observed between the exercise bouts. Accordingly, GXT was found to be valid when VER $\dot{V}O_{2max}$ was less than 3% higher that observed during GXT (25, 28), following previous recommendations (16, 17). Finally, the agreement between GXT and VER $\dot{V}O_{2max}$ measures was assessed using Bland-Altman plots and reliability was measured using the coefficient of variation. The level of significance was set at $\alpha = 0.05$ and all analyses were performed using SPSS software (v. 25.0., IBM, USA).

RESULTS

Mean comparisons

Testing was well tolerated by all participants, and no adverse event was observed. Mean data from both GXT and VER are presented in Table 2. Both absolute and relative $\dot{V}O_{2max}$ values were higher in VER than in GXT ($p = 0.011$ and 0.014, respectively). RER and RPE were lower in VER ($p < 0.001$ and $p = 0.002$, respectively), whereas HR_{max} was not different between the two trials ($p = 0.286$). Coefficient of variation for $\dot{V}O_{2max}$ was 8.1% (CI95%: 6.1 to 10.1%) between GXT and VER. Bland-Altman plots are presented in Figure 1. Accordingly, no agreement was observed for both absolute ($p = 0.011$) and relative ($p = 0.014$) $\dot{V}O_{2max}$ results between trials, and there was no evidence of proportional bias ($p = 0.300$, $p = 0.189$, respectively).

Insert Table 2 near here

Insert Figure 1 near here

Individual comparisons

$\dot{V}O_{2max}$ comparison between GXT and VER based on individual tests revealed that 54.6% of the participants (18/33) achieved a $\dot{V}O_{2max}$ value that was ≥3% higher during VER (range: from 3% to 22.1%). A total of 24.2% (8/33) had VER $\dot{V}O_{2max}$ within ±2.9% of their GXT result (range: from -2.9 to 1.5%), whereas 21.2% (7/33) displayed VER results that were ≥3% lower than GXT (range: from -5.1 to -29.4%). A waterfall plot of the individual differences between VER and GXT is shown in Figure 2.

Insert Figure 2 near here

Classic criteria

When GXTs were analyzed based on the classic established criteria (i.e. $\dot{V}O_2$ plateau and/or secondary criteria), results diverged from those observed in the individual comparisons between VER and GXT (Table 3). Specifically, if only one criterion was used to validate GXT, 97% (32/33) of the participants would have had their tests confirmed as valid. If at least two criteria were used, as proposed by Howley et al. (12), 87.9% (29/33) of the present sample would have had their GXT validated. When only those tests that were deemed to be underestimated during GXT were analyzed separately, 94.4% (17/18) of them would have been confirmed as valid based on the occurrence of a $\dot{V}O_2$ plateau or the presence of at least two secondary criteria.

Insert Table 3 near here

As an additional analysis, GXTs that would have been validated by classic criteria were further separated according to the specific criteria that were reached and then compared to VER results (see Table, Supplemental Digital Content, Comparison between GXT and VER $\dot{V}O_{2max}$ results stratified according to the primary and secondary criteria achieved during GXT). The analyses revealed that both absolute and relative $\dot{V}O_{2max}$ were lower in GXT for those who achieved (i) a plateau ($\leq$ 150 $ml \cdot min^{-1}$; $p < 0.001$); (ii) HR_{max} and RER ($p \leq 0.022$); (iii) HR_{max} and RPE ($p \leq 0.049$); (iv) as well as RER and RPE ($p \leq 0.004$) criteria. No difference was observed between GXT and VER $\dot{V}O_{2max}$ for those who presented a plateau as proposed by Midgley et al. ($p = 0.063$) or who reached all three secondary criteria ($p = 0.075$ and 0.059). In addition, RER and RPE were lower during VER ($p \leq 0.034$) and no difference was found in HR_{max} between VER and GXT ($p > 0.05$).

DISCUSSION

The main finding of the present study was that in hypertensive older individuals GXT underestimated $\dot{V}O_{2max}$ in over half of the participants, as demonstrated by VER, even though close to ninety percent of the tests would have been validated as a maximal effort if the classic criteria were employed. Moreover, VER testing appeared to be safe for the study population with no adverse event registered. Hence, our data demonstrate that VER can be a practical and sensitive method to verify GXT $\dot{V}O_{2max}$ and can be performed during the same session as GXT in hypertensive individuals.

This is the first study to our knowledge that investigated VER in older adults with hypertension. Most of the previous studies were performed on healthy young men and women and showed that

VER is indeed able to confirm whether $\dot{V}O_{2max}$ was attained (25, 28–32). Our results, however, suggest otherwise in the present population, in disagreement with previous data on physically active middle-aged and older adults (18). Using a constant-load VER at 105% of the peak power output, Dalleck et al. (18) found no difference between GXT and VER mean $\dot{V}O_{2max}$ results, whereas only 11% of their older participants reached $\dot{V}O_2$ values during VER that were greater than GXT by more than 3%. This contrasts with our current findings, with 55% of participants displaying greater $\dot{V}O_2$ during VER compared to GXT. In fact, when only the underestimated tests were considered, $\dot{V}O_{2max}$ was found to be, on average, 13.5% (~3.5 $ml \cdot kg^{-1} \cdot min^{-1}$) higher in VER (Cohen's d = 0.62). Such marked differences between studies may be related to several factors, including participant's motivation to achieve maximal effort and the protocols used during the VER phase.

An important aspect that might have influenced the aforementioned results is the difference in fitness between our participants and that of Dalleck et al.'s. The latter was composed mainly of older individuals who were actively engaged in an exercise program. Supporting this notion, it was previously demonstrated that while no differences were observed between GXT and VER among young individuals with moderate to high fitness levels ($\dot{V}O_{2max}$ > than 40 and 50 $ml \cdot kg^{-1} \cdot min^{-1}$, respectively), those considered as low fitness ($\dot{V}O_{2max}$ > 30 $ml \cdot kg^{-1} \cdot min^{-1}$) achieved higher $\dot{V}O_{2max}$ during VER (33). This is also in line with studies on cancer (34) and chronic heart failure patients (35), which are conditions typically related to lower exercise tolerance and fitness level. Indeed, most of the literature investigating VER ability to confirm $\dot{V}O_{2max}$ on sedentary individuals was performed on inactive but otherwise healthy men and women, with relatively low cardiovascular risk (1, 36, 37). Thus, those with chronic conditions, such as hypertension, have been overlooked.

Because mean value comparisons might mask a considerable between-trials variability in some participants, previous authors recommended that comparisons between GXT and VER $\dot{V}O_{2max}$ should be performed on an individual case-by-case basis (15, 16, 38–40). Our results demonstrated not only a substantial $\dot{V}O_{2max}$ underestimation when individual data were compared, but also an underestimation when mean values were used for analysis, which might relate to the multistage VER protocol approach employed in this study.

Square-wave tests are the most common VER protocols employed in previous investigations (16). In such studies, the entire VER is performed at a constant load that typically can be sustained for ~2 min (e.g., 18, 30, 33, 35, 37, 41). When a multistage protocol is performed, however, participants are allowed a certain time for $\dot{V}O_2$ to adjust before the final workload is delivered. In the case of the present study, this approach resulted in VER duration of about 5 min on average. Even though the duration of exercise at supra GXT intensity might be similar or slightly lower than those of square-wave protocols, it is possible that the first submaximal stages allowed participants to reach the final stage at a higher $\dot{V}O_2$ and, in turn, less time was possibly required at this supramaximal intensity to reach or surpass the $\dot{V}O_2$ observed during GXT.

It has been reported that a previous heavy intensity exercise bout may speed $\dot{V}O_2$ kinetics while attenuating the $\dot{V}CO_2$ response during a subsequent bout of heavy or maximal exercise, leading to reduced RER values, as observed in VER compared to GXT in the present study (42, 43). Considering that less fit individuals normally possess a lower anaerobic capacity and slower $\dot{V}O_2$ kinetics, multistage approaches may also reduce the anaerobic burden and assist these participants

to sustain efforts for longer (1, 26). Although not yet tested, this hypothesis is supported by the lower RPE and RER reported after VER compared to GXT.

In accordance with our initial hypothesis, 85% of the participants would have had their maximal tests considered valid based on the achievement of at least two of the secondary criteria investigated (Table 3), even though VER demonstrated a considerable $\dot{V}O_{2max}$ underestimation during GXT. As such, the use of these secondary criteria has been questioned (14, 17, 25, 44), with some authors reporting that the same secondary criteria employed in the current study (e.g. HR_{max} and RER) may be reached as early as 80% $\dot{V}O_{2max}$. These criteria are also affected by the rate at which GXT workload is increased. For example, final RER has been reported to be higher when the workload is increased at a faster rate (35). Accordingly, these criteria have been strongly criticized for their validity and sensitivity to verify $\dot{V}O_{2max}$ (14, 15). This is further reinforced by the differences between GXT and VER that were maintained even when data were stratified between primary and secondary criteria (see Table, Supplemental Digital Content, Comparison between GXT and VER $\dot{V}$O2max results stratified according to the primary and secondary criteria achieved during GXT).

This additional exploratory analysis indicated that $\dot{V}O_{2max}$ was not different when GXT and VER were compared only in those who achieved either a $\dot{V}O_2$ plateau as proposed by Midgley or all three secondary criteria (see Table, Supplemental Digital Content, Comparison between GXT and VER $\dot{V}$O2max results stratified according to the primary and secondary criteria achieved during GXT). Such results would indicate that a plateau determined based on the individual $\dot{V}O_2$ rate of increase, rather than a fixed or average increase (e.g. $< 150\ ml \cdot min^{-1}$), would likely be more suitable

to validate and confirm $\dot{V}O_{2max}$ attainment (25), whereas participants reaching both HR_{max}, RPE and RER criteria would have provided a greater effort during the incremental test compared to those who have reached only two. In this regard, however, it is important to note that likely due to their more stringent nature these criteria were the ones fulfilled by the least number of participants (n = 11 and 12, respectively). It is possible, therefore, that these exploratory analyses lacked the power to detect such differences. Irrespective of this, the aforementioned results suggest that, if a criterion is to be found to validate GXT on its own, an individualized approach is more likely to succeed (25).

LIMITATIONS

The present study is not exempt from limitations. First, the 3% cutoff threshold adopted here to validate a difference between GXT and VER was not based on data from our laboratory and, therefore, it is possible that the GXT test-retest reliability in our laboratory might be slightly different than the one used. Nevertheless, this value has been suggested previously (28) and employed with some degree of success by others in young (11, 30, 32, 45), middle-aged, and older adults (18). In fact, even if a 9% cutoff value was adopted, as previously reported in cystic fibrosis patients, more than 40% of the current tests would still result in higher $\dot{V}O_{2max}$ during VER. In addition, it should be noted that the lack of a non-hypertensive control group is a limitation to make inferences about the influence of hypertension *per se* on our findings and additional research is necessary to elucidate this. Finally, the absence of a sample size estimation may also be suggested as a limitation, however, our study seemed enough powered (all $\geq$ 0.82) to detect differences in all outcomes that were expected to differ between GXT and VER), as shown in Table 2.

CONCLUSION

In conclusion, our results demonstrate the relevance that verifying the $\dot{V}O_{2max}$ obtained during GXT with a relatively simple strategy can have in clinical populations such as the one studied in the present investigation. The use of VER might be an alternative approach to consider, as it is not very time-consuming and enables direct $\dot{V}O_2$ comparisons between test phases. Thus, even though it is not possible to assume that VER represents the participant's "true" $\dot{V}O_{2max}$ when it is higher than GXT, by implementing a VER researchers are provided with the highest $\dot{V}O_2$ from the two phases. This would not be possible through other criteria based solely on a single incremental test (e.g., HR_{max} or RER). This is especially important when the performance of another GXT is not feasible because of logistics, such as during large clinical trials.

Acknowledgments

The HAEL Study or involved personnel have been funded by CNPq, grant number 429849/2016-8 (Conselho Nacional de Desenvolvimento Científico e Tecnológico), FAPERGS (Fundação de Amparo à Pesquisa do Estado do Rio Grande do Sul), CAPES (Coordenação de Aperfeiçoamento de Pessoal de Nível Superior), and FIPE (Fundo de Incentivo à Pesquisa e Eventos, Hospital de Clinicas de Porto Alegre). The authors are thankful for the administrative and funding support provided by the National Institute of Science and Technology for Health Technology Assessment (IATS/HCPA), Hospital de Clínicas de Porto Alegre, Porto Alegre, RS, Brazil. E. N. W. was supported by the Brazilian Education Ministry Foundation CAPES (PNPD).

Data availability

Data from the present study will be made available upon request to the corresponding author.

Conflicts of Interest

All authors declare that they have no competing financial interests related to this publication. All authors declare that the results of the study are presented clearly, honestly, and without fabrication, falsification, or inappropriate data manipulation. The results of the present study do not constitute an endorsement by the American College of Sports Medicine.

REFERENCES

1. Astorino TA, White AC, Dalleck LC. Supramaximal testing to confirm attainment of VO2max in sedentary men and women. *Int J Sports Med*. 2009;30(4):279–84.
2. Hill A V, Lupton H. Muscular Exercise, Lactic Acid, and the Supply and Utilization of Oxygen. *QJM An Int J Med*. 1923;os-16(62):135–71.
3. Day JR, Rossiter HB, Coats EM, Skasick A, Whipp BJ. The maximally attainable V̇O2 during exercise in humans: the peak vs. maximum issue. *J Appl Physiol*. 2003;95(5):1901–7.
4. Balady GJ. Survival of the Fittest — More Evidence. *N Engl J Med*. 2002;346(11):852–4.
5. Kodama S, Saito K, Tanaka S, et al. Cardiorespiratory Fitness as a Quantitative Predictor

of All-Cause Mortality and Cardiovascular Events in Healthy Men and Women: A Meta-analysis. *JAMA*. 2009;301(19):2024–35.

6. 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.

7. Astorino TA, Robergs RA, Ghiasvand F, Marks D, Burns S. Incidence Of The Oxygen Plateau at VO2max During Exercise Testing To Volitional Fatigue. *J Exerc Physiol online*. 2000;3(4):1–12.

8. Ferretti G. Maximal oxygen consumption in healthy humans: theories and facts. *Eur J Appl Physiol*. 2014;114(10):2007–36.

9. Howley ET. VO2max and the Plateau - Needed or Not? [Internet]. *Med Sci Sport Exerc*. 2007;39(1) available from: https://journals.lww.com/acsm-msse/Fulltext/2007/01000/V_O2max_and_the_Plateau_Needed_or_Not_.16.aspx.

10. Midgley AW, Mcnaughton L. Criteria for Determination of Maximal Oxygen Uptake: A Brief Critique and Recommendations for Future Research. *Sport Med*. 2007;12(37):1019–28.

11. Sedgeman D, Dalleck L, Clark IE, Jamnick N, Pettitt RW. Analysis of square-wave bouts to verify VO2max. *Int J Sports Med*. 2013;34(12):1058–62.

12. Howley ET, Bassett DRJ, Welch HG. Criteria for maximal oxygen uptake: review and commentary. *Med Sci Sports Exerc*. 1995;27(9):1292–301.

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

14. Beltz NM, Gibson AL, Janot JM, Kravitz L, Mermier CM, Dalleck LC. Graded Exercise Testing Protocols for the Determination of VO2max: Historical Perspectives, Progress,

and Future Considerations. *J Sports Med*. 2016;2016:1–12.

15. Midgley AW, Carroll S. Emergence of the verification phase procedure for confirming "true" VO2max. *Scand J Med Sci Sport*. 2009;19(3):313–22.

16. Schaun GZ. The Maximal Oxygen Uptake Verification Phase: a Light at the End of the Tunnel? *Sport Med - Open*. 2017;3(44):1–15.

17. Poole DC, Jones AM. Measurement of the maximum oxygen uptake V̇O2max: V̇O2peak is no longer acceptable. *J Appl Physiol*. 2017;122(4):997–1002.

18. Dalleck LC, Astorino TA, Erickson RM, McCarthy CM, Beadell AA, Botten BH. Suitability of verification testing to confirm attainment of VO2max in middle-aged and older adults. *Res Sport Med*. 2012;20(2):118–28.

19. Mills KT, Bundy JD, Tanika KN, et al. Global Disparities of Hypertension Prevalence and Control. *Circulation*. 2016;134(6):441–50.

20. American College of Sports Medicine. *ACSM's Guidelines for Exercise Testing and Prescription*. 10th ed. Philadelphia: Wolters Kluwer Health/Lippincott Williams & Wilkins; 2017. 472 p.

21. Thomas SG, Goodman JM, Burr JF. Evidence-based risk assessment and recommendations for physical activity clearance: Established cardiovascular disease. *Appl Physiol Nutr Metab*. 2011;36(SUPPL.1):190–213.

22. Fletcher GF, Ades PA, Kligfield P, et al. Exercise standards for testing and training: A scientific statement from the American Heart Association. *Circulation*. 2013;128(8):873–934.

23. Ostchega Y, Fryar CD, Nwankwo T, Nguyen DT. Hypertension Prevalence Among Adults Aged 18 and Over: United States, 2017-2018. *NCHS Data Brief*. 2020;(364):1–8.

24. Umpierre D, Santos LP, Botton CE, et al. The “Hypertension Approaches in the Elderly: a Lifestyle study” multicenter, randomized trial (HAEL Study): rationale and methodological protocol. *BMC Public Health*. 2019;19(1):657.

25. Midgley AW, Carroll S, Marchant D, McNaughton LR, Siegler J. Evaluation of true maximal oxygen uptake based on a novel set of standardized criteria. *Appl Physiol Nutr Metab*. 2009;34(2):115–23.

26. Poole DC, Jones AM. Oxygen Uptake Kinetics. *Compr Physiol*. 2012;933–96. Available from: https://doi.org/10.1002/cphy.c100072.

27. Robergs RA, Dwyer D, Astorino T. Recommendations for Improved Data Processing from Expired Gas Analysis Indirect Calorimetry. *Sport Med*. 2010;40(2):95–111.

28. Kirkeberg JM, Dalleck LC, Kamphoff CS, Pettitt RW. Validity of 3 protocols for verifying VO2max. *Int J Sports Med*. 2011;32(4):266–70.

29. Foster C, Kuffel E, Bradley N, et al. VO2max during successive maximal efforts. *Eur J Appl Physiol*. 2007;102(1):67–72.

30. Astorino TA, White AC. Assessment of anaerobic power to verify VO2max attainment. *Clin Physiol Funct Imaging*. 2010;30(4):294–300.

31. 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.

32. Nolan PB, Beaven ML, Dalleck L. Comparison of intensities and rest periods for VO2max verification testing procedures. *Int J Sports Med*. 2014;35(12):1024–9.

33. Astorino TA, DeRevere J. Efficacy of constant load verification testing to confirm VO2max attainment. *Clin Physiol Funct Imaging*. 2018;38(4):703–9.

34. Schneider J, Schlüter K, Wiskemann J, Rosenberger F. Do we underestimate maximal oxygen uptake in cancer survivors? Findings from a supramaximal verification test. *Appl Physiol Nutr Metab*. 2020;45(5):486–92.

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

36. Wood RE, Hills AP, Hunter GR, King NA, Byrne NM. V̇O2max in overweight and obese adults: Do they meet the threshold criteria? *Med Sci Sports Exerc*. 2010;42(3):470–7.

37. Sawyer BJ, Tucker WJ, Bhammar DM, Gaesser GA. Using a Verification Test for Determination of VO2max in Sedentary Adults With Obesisty. *J Strength Cond Res*. 2015;29(12):3432–8.

38. Pettitt RW, Clark IE, Ebner SM, Sedgeman DT, Murray SR. Gas Exchange Threshold and VO2max Testing for Athletes. *J Strength Cond Res*. 2013;27(2):549–55.

39. Hecksteden A, Kraushaar J, Scharhag-Rosenberger F, Theisen D, Senn S, Meyer T. Individual response to exercise training - a statistical perspective. *J Appl Physiol*. 2015;118(12):1450–9.

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

41. Moreno-Cabañas A, Ortega JF, Morales-Palomo F, et al. The use of graded exercise test may be insufficient to quantify true changes in VO2max following exercise training in unfit individuals with metabolic syndrome [Internet]. *J Appl Physiol*. 2020; available from: https://doi.org/10.1152/japplphysiol.00455.2020. doi:10.1152/japplphysiol.00455.2020.

42. Gerbino A, Ward SA, Whipp BJ. Effect of prior exercise on pulmonary gas-exchange kinetics during high-intensity exercse in humans. *J Appl Physiol*. 1996;80(1):99–107.

43. Jones AM, Wilkerson DP, Burnley M, Koppo K. Prior Heavy Exercise Enhances Performance during Subsequent Perimaximal Exercise. *Med Sci Sports Exerc*. 2003;35(12):2085–92.

44. 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.

45. Weatherwax RM, Richardson TB, Beltz NM, Nolan PB, Dalleck L. Verification Testing to Confirm VO2max in Altitude- Residing, Endurance-Trained Runners. 2015;24:3–8.

Figure legends

Figure 1. Bland-Altman plots for the absolute (A) and relative (B) maximal oxygen uptake ($\dot{V}O_{2max}$) responses between the graded exercise test and verification phase. The solid line represents the mean bias and dashed lines the limits of agreement (mean ±1.96*SD).

Figure 2. Mean (bars) and individual (lines) maximal oxygen uptake ($\dot{V}O_{2max}$) responses between the graded exercise test (GXT) and the verification phase (VER) (A). Waterfall plot (B) of the individual differences in $\dot{V}O_{2max}$ ($\Delta\dot{V}O_{2max}$) between VER and GXT. In (B), values within the dotted lines (gray bars) represent participants with a lower than 3% difference between VER and GXT $\dot{V}O_{2max}$ results, whereas those above (black bars) and below (white bars) correspond to differences greater and lower than 3%, respectively (as proposed by Midgley et al. 2009). * = significantly different from GXT (p = 0.014).

Supplemental Digital Content

Supplementary Table 1. Comparison between GXT and VER $\dot{V}O_{2max}$ results stratified according to the primary and secondary criteria achieved during GXT

Table 1.

Table 1. Participants' characteristics (n = 33).

	Mean ± SD
Age (years)	67.1 ± 5.1
Sex	24 female / 9 male
Race	24 white / 9 black
Height (m)	1.55 ± 0.09
Body mass (kg)	77.3 ± 13.8
Waist circumference (cm)	103.2 ± 10.6
SBP (mmHg)	128 ± 10.8
DBP (mmHg)	72 ± 7.4
Antihypertensive medications*	
Diuretics (n)	19
ARB (n)	23
CCB (n)	5
Beta-blockers (n)	11
ACE inhibitors (n)	8

SBP 24-h ambulatory systolic blood pressure; *DBP 24-h ambulatory diastolic blood pressure*. *ARB angiotensin receptor blockers; CCB calcium channel blockers ACE angiotensinogen converting enzyme*.
*Number of medications per drug class is presented as frequency.

Table 2.

Table 2. GXT and VER outcomes in older adults with hypertension (mean ± SD).

	Incremental phase	Verification phase	p
$\dot{V}O_{2max}$ ($l \cdot min^{-1}$)	1.72 ± 0.46	1.83 ± 0.50	**0.011**
$\dot{V}O_{2max}$ ($ml \cdot kg^{-1} \cdot min^{-1}$)	22.5 ± 4.9	23.9 ± 5.6	**0.014**
HR_{max} ($beats \cdot min^{-1}$)	150.4 ± 16.4	151.8 ± 15.7	0.286
RER	1.15 ± 0.18	0.97 ± 0.08	**<0.001**
RPE	17.7 ± 1.9	16.7 ± 2.3	**0.002**
Duration (s)	714.8 ± 126.8	286.5 ± 26.5	**<0.001**

$\dot{V}O_{2max}$ maximal oxygen uptake; HR_{max} maximal heart rate; *RER* respiratory exchange ratio; *RPE* rating of perceived exertion.

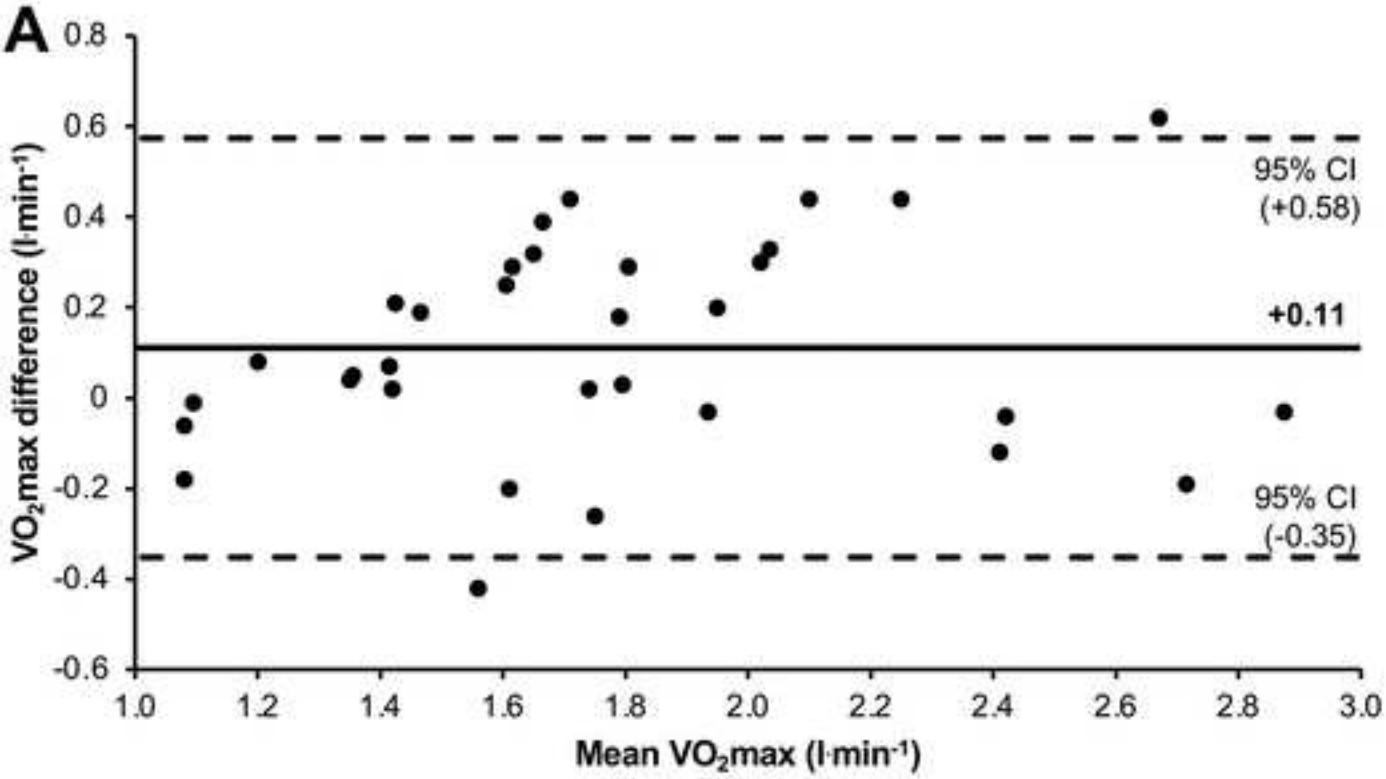
Table 3.

Table 3. Frequency of primary and secondary criteria attainment during the incremental test and number of tests deemed valid (n = 33).

	Frequency (%)	
	No	Yes
$\dot{V}O_2$ plateau (Midgley et al.)	23 (69.7)	10 (30.3)
$\dot{V}O_2$ plateau (Taylor et al.)	18 (54.5)	15 (45.5)
HR_{max}	9 (27.3)	24 (72.7)
RER	12 (36.4)	21 (63.6)
RPE	9 (27.3)	24 (72.7)
*Validated tests**		
One criterion	1 (2.9)	32 (96.7)
Two criteria	5 (15.1)	28 (84.9)

$\dot{V}O_{2max}$ maximal oxygen uptake; HR_{max} maximal heart rate; *RER respiratory exchange ratio;* RPE rating of perceived exertion. *results are similar using both plateau criterion.

Figure 1.



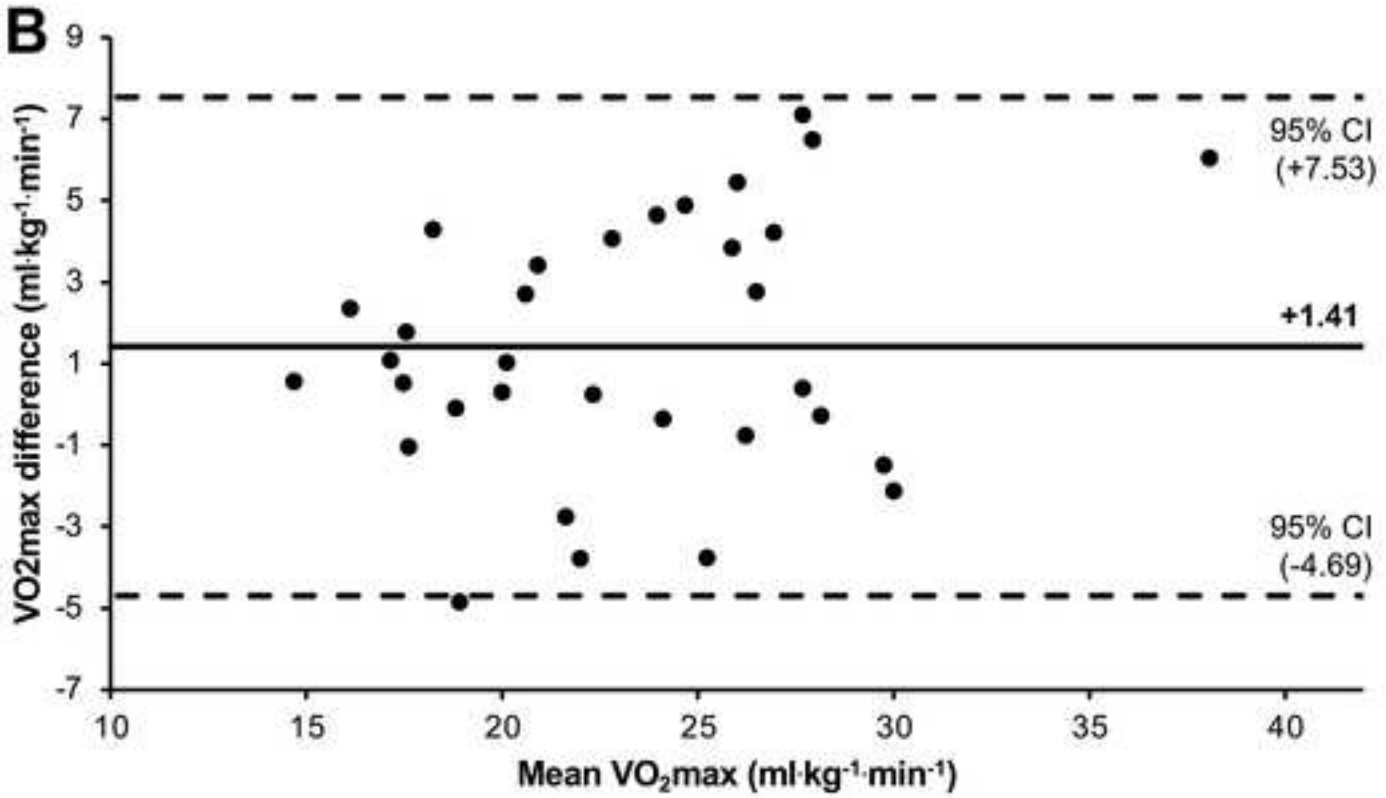
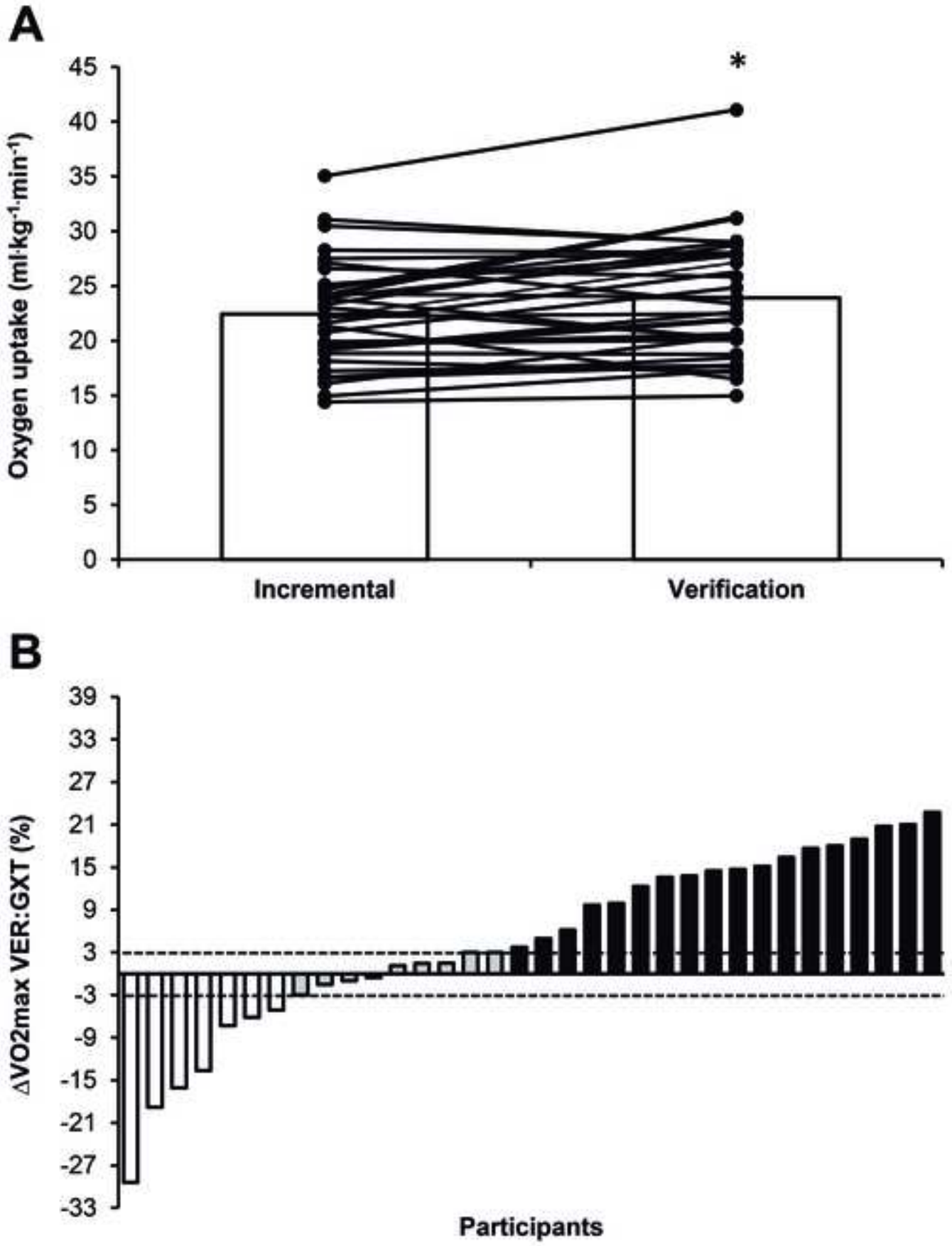


Figure 2.



Supplemental Table 1.