5-8-2020

The efficacy of a verification stage for determining VO2max and the impact of sampling time

Emily J. Kontos

The Efficacy of a Verification Stage for Determining VO_{2max} and the Impact of Sampling Time

An Honors College Project Presented to

The Faculty of the Undergraduate

College of Health and Behavioral Science

Department of Kinesiology

James Madison University

Emily J. Kontos

April 30, 2020

Accepted by the faculty of the Kinesiology Department, James Madison University, in partial fulfillment of the requirements for the Honors College.

FACULTY COMMITTEE:

Project Advisor: Christopher J. Womack, Ph.D., FACSM
Professor, James Madison University

Reader: Stephanie P. Kurti, Ph.D.
Assistant Professor, James Madison University

Reader: Nicholas D. Luden, Ph.D.
Professor, James Madison University

HONORS COLLEGE APPROVAL:

Bradley R. Newcomer, Ph.D.,
Dean, Honors College

PUBLIC PRESENTATION

This work was accepted for presentation and displayed at the 2020 ACSM Southeast Conference in Jacksonville, Florida, 15th of February, 2020.

Dedication

I dedicate this thesis to all of the amazing friends, students, and professors I have had the privilege to work with in the Kinesiology Department over the past four years at James Madison University.

Acknowledgments

Firstly, I would like to thank Dr. Christopher J. Womack, my thesis advisor and mentor, for his constant support, enthusiasm, and patience over the past three semesters of this project. Without his knowledge, expertise, and passion for kinesiology, I'm not sure if this project would have come to fruition. Through working with him, I have come to develop a greater appreciation of research and have found my own passion for exercise physiology. I will forever be thankful and appreciative of his time and efforts as both a mentor and a professor.

Secondly, I would like to thank both Dr. Stephanie Kurti and Dr. Nick Luden for agreeing to be my readers for this project. Their added advice, expertise, and assistance are greatly appreciated. Without their assistance, this project wouldn't be what it is.

Thirdly, I would like to thank my mother and father for their continued support and their encouragement. My mother's enthusiasm and dedication for women in science helped provide fuel in moments of doubt. My father's assistance in reading and answering numerous cardiac questions helped me better understand my own research. Without them, I wouldn't have enjoyed this project nearly as much.

Finally, I would like to thank my fellow James Madison University kinesiology students and friends. Not only did several friends assist me in running VO_{2max} tests, they supported me and rooted me on, even when I felt like this project would never be finished. I will forever be grateful for the support of my peers.

Table of Contents

List of Tables and Figures

Abstract

Purpose: Verification phases may improve the validity of maximal oxygen uptake (VO_{2max}) measurements during maximal graded exercise testing (GXT). It is not known whether VO_2 sampling times influence the necessity of a verification stage. **Methods:** 15 female and 18 male test subjects (18 – 25 y) completed a treadmill incremental GXT. Speed was increased from 3.0 mph by 0.5 mph every minute until 6.0 mph was reached. Elevation was then increased by 3% every minute until volitional fatigue. Subjects then walked for five minutes at 3.0 mph and 0% grade; after which time the verification stage began at the speed and grade corresponding with the penultimate stage and continued until volitional fatigue. VO_{2max} from the incremental GXT (iVO_{2max}) and VO_{2max} from the verification stage ($verVO_{2max}$) were determined using 10 s, 30 s and 60 s averages from the breath x breath measurements. A repeated measures ANOVA was performed with sampling time (10, 30 and 60s) and stage (iVO_{2max}, $verVO_{2max}$) as the within-subject factors. Sensitivity and specificity were calculated for the following criteria from the iVO_{2max} portion of the protocol: plateau (< 150 ml/min increase in VO_2 over the final 2 stages), and HR + RER (achievement of at least 90% of age-predicted maximal heart rate and RER $\geq$ 1.10). **Results:** There was no main effect for stage, suggesting no differences between iVO_{2max} and $verVO_{2max}$ for 10s (47.9 ± 8.31 ml/kg/min vs 48.85 ± 7.97 ml/kg/min), 30s (46.94 ± 8.62 ml/kg/min vs 47.28 ± 7.97 ml/kg/min), and 60s (46.17 ± 8.62 ml/kg/min vs 46.00 ± 8.00 ml/kg/min) sampling times. There was a main effect for sampling time for VO_{2max} (10s > 30s > 60s, $P < 0.05$). Furthermore, there was a significant ($P < 0.05$) stage x sampling time interaction as the difference between iVO_{2max} and $verVO_{2max}$ was greater for 10s than 60s sampling times. $verVO_{2max}$ was considered to be higher if it exceeded iVO_{2max} by more than 2%, as suggested by Midgley et al[10]. This was seen in 62%, 41%, and 31% of the tests for the 10s, 30s and 60s sampling times respectively. Both sensitivity and specificity for the plateau criteria was under 45% for all sampling times. Sensitivity of using HR + RER was above 80% for all sampling times and specificity was under 30%.

Conclusions: A verification stage yields a higher VO_{2max} in a large proportion of tests and the effectiveness of the verification stage may be more important with shorter sampling times. A plateau for determining the achievement of VO_{2max} during an incremental test has poor sensitivity and specificity and the use of HR + RER criteria exhibits poor specificity.

Keywords: VO_{2max} testing, verification stage, VO_{2max} criteria specificity, VO_{2max} criteria sensitivity, sampling time

Chapter I.

Introduction

For decades, cardiorespiratory fitness has been determined through maximal oxygen consumption (VO_{2max}) testing. VO_{2max}, determined via graded exercise tests, has been associated with long-term health outcomes[19], including mortality, cardiovascular disease risk[26], and other pathological conditions that impact the pulmonary, cardiovascular, and muscular systems[21]. VO_{2max} values have also been linked to incidences of chronic heart disease, diabetes, and HIV-AIDS[21]. Ever since it's development in 1923[27], the VO_{2max} graded exercise test (GXT) has been utilized often and is a commonly measured variable. However, despite its widespread use, there still lacks a standardization for criteria used to verify that test subjects have reached their true VO_{2max}.

Since its development by Hill and Lupton[9], VO_{2max} testing protocols have varied in stage length, stage variability, and test length[27]. Traditionally, a plateau in VO_2 at the end of a GXT has been the primary criteria for confirming the achievement of VO_{2max}. A plateau is generally defined as a period at the end of the test in which there is little to no increase in VO_2 despite an increased work rate. However, the specific criteria vary from as little as an increase <54 mL/min to <2.1 mL/kg/min[10]. The most common VO_{2max} plateau criterion stipulates an increase < 150 mL/min, which was developed in 1955 by Taylor et al[10]. However, the effectiveness of this value may be specific to the test protocol and the subject pool size used[10]. Additionally, unless an absolute plateau (i.e. no increase in VO_2 whatsoever) is employed, utilization of other criteria will only show that the rate of change between VO_2 and work-rate has slowed. It will not be an indication that a VO_2max, per se, has been achieved[16]. The incidence of a plateau has been reported as low as 33% and has high as 94% in subjects, depending on the criterion used, the test protocol, and the test subjects[27]. Other literature reports that a plateau occurs ≥ 40% of healthy test subjects, and likely occurs less often in clinical populations[21].

Therefore, in the absence of the occurrence of a VO_{2max} plateau, other secondary criteria have been used in order to verify that VO_{2max} was achieved.

Currently there is no standardization of secondary criteria. As a result, both the variables and the critical value for those variables vary[16]. Common secondary criteria include a heart rate ≥ 10 bpm or ≤ 5% of age-predicted maximal heart rate (220 bpm - age of test subject), a blood lactate concentration ≥ 8 mM, and/or an RER (respiratory exchange rate) between 1.0-1.44[21]. However, these criteria are selected arbitrarily and without the support of research[21]. Several studies have found them to be unreliable, as many of the test subjects will fulfill one or more of the secondary criteria during submaximal tests[16, 17]. As a result, implementing these criteria can result in an underestimation of VO_{2max} data by 30-40%[21].Furthermore, testing protocols often vary with respect to stage length, magnitude of work rate increase during each stage, and test duration[27], thus making the efficacy of primary and secondary criteria across all protocols questionable[16].

Therefore, improvement in the methodology of VO_{2max} testing and/or criteria that are used are warranted. Recently, verification stages have been suggested as a method to improve the validity of VO_{2max} testing. These verification stages are an exercise stage performed after the incremental, maximal test to exhaustion. Their intensities range from submaximal efforts to supramaximal efforts following a recovery period lasting between five and 15 minutes[24]. Currently, verification stages are used infrequently, and the specific details of what an appropriate verification stage entails have not been fully evaluated. Furthermore, it is not definitively known if the verification stage is noticeably better than using traditional criteria.

Several studies have found that when a subject completes a VO_{2max} test, a subsequent verification stage results in subjects either reaching or exceeding the stage in which they originally stopped the test[13, 17, 7]. Foster et al., had test subjects perform both a cycling GXT and a treadmill GXT. Once they achieved volitional exhaustion during the continuous portion of the GXT, they were given a

rest period followed by a supramaximal verification stage. This study found that the VO_2 values achieved during the verification stage was not significantly different than the values observed during the continuous GXT for both running and cycling. Furthermore, during verification stages, subjects attain HR and RER values comparable or higher than during the original GXT test[17, 7].Thus, Midgley et al.[16] have argued that use of a verification stage increases the probability that VO_{2max} is achieved during the test, as it provides an additional opportunity to reach the limits of oxygen uptake. Practically speaking a verification stage will either verify attainment of VO_{2max} during the GXT or result in a higher measurement for VO_{2max}.

Mier and colleagues had 35 male and female college athletes complete a treadmill GXT test. Ten subjects who did not exhibit a VO_{2max} plateau during the continuous GXT completed a supramaximal verification stage. VO_{2max} values from the verification stage were not significantly different than VO_{2max} values from the GXT. Additionally, four out of the 10 subjects achieved a higher VO_{2max} in the verification stage. These findings support the effectiveness of a verification stage to verify or correct VO_{2max} measurements from a GXT. They also illustrate that a means comparison between VO_{2max} values from the GXT and the verification stage is not enough to fully determine the utility of a verification stage. The fact that 40% of the subjects did not achieve VO_{2max} in the GXT would likely be concerning for any lab, but this is not illuminated by simply showing a lack of difference in mean values. It was also suggested that HR and RER criteria are ineffective in verifying whether a VO_{2max} was achieved, as common secondary criteria were often achieved at a submaximal effort[17.]

Another investigation conducted by Foster and colleagues had physically active, non-athletes complete a cycling incremental exercise test. After the test, a one-minute recovery phase was performed, followed by a verification stage at a higher power output than achieved during the GXT. Only subjects exhibiting a plateau during the GXT were included in the data analysis to ensure that

VO_{2max} was actually achieved during that portion of the test. The verification stage resulted in the same average VO_{2max} values, suggesting that true VO_{2max} values are achieved during a verification stage[7]. This same study observed similar findings in runners performing a treadmill GXT[7].

Higher VO_{2max} plateau incidences have been observed for 11 and 15 seconds compared to 30 second sampling averages for breath by breath VO_2 measurements[2]. Furthermore, 15 and 30s intervals have shown to result in higher VO_{2max} values than 60s intervals[1, 23]. Therefore, the need for a verification stage may depend upon the duration of the sampling time. It is not known if use of different sampling times impacts the effectiveness of traditional primary and secondary criteria for determining VO_{2max} from a GXT. The use of a verification stage to evaluate the relative sensitivity and/or specificity of common criteria used for VO_{2max} testing is limited in the literature. Bhammar et al. found that there was a low sensitivity and low specificity for traditional VO_2 criteria. However, this study was performed with a limited number of test subjects and the test subjects were children. Thus, these findings are limited in their generalizability[4]. The purposes of this present study were to: 1) determine the influence of sampling time on the efficacy of a verification stage; and 2) to determine the sensitivity and specificity of primary and secondary VO_{2max} test criteria. It is hypothesized that a verification stage VO_{2max} will exceed the VO_{2max} reached in an incremental, graded exercise test in proportion to sampling time and that traditional primary and secondary criteria for achievement of VO_{2max} will exhibit poor sensitivity and specificity regardless of sampling time.

Chapter II.

Methods

Subjects: This study will evaluate 29 test subjects (14 men, 15 women). All test subjects will be between 18-30 years of age upon the day of testing (men = 21.5 ± 1.2 yr, women = 21.1 ±1.2 yr). Before testing, subjects will complete a health questionnaire and will have no known CV, metabolic, or renal disease. Additionally, test subjects will have no known injuries or other health concerns that would preclude them from exercise or limit their ability to perform a maximal exercise bout.

Treadmill Test: All subjects will be monitored for oxygen uptake (VO_2) and respiratory exchange ratio (RER) with a Vmax metabolic cart (CareFusion; San Diego, CA) throughout the duration of the test. A Polar heart rate monitor will be utilized to measure heart rate. After an initial stage at 3.0 mph and 0% grade, the treadmill speed will be increased by 0.5 mph each minute until a speed of 6.0 mph is achieved. After this, the elevation will be increased by 3% every minute until volitional exhaustion. Subjects will then be allowed to walk for 5 minutes at 3.0 mph and 0% grade. After this rest period, the verification stage will be initiated by increasing the speed and grade to the values of the stage preceding the test subject's prior maximal effort. The test will then proceed as described previously until volitional exhaustion.

Statistical Analyses: A repeated measure of analysis of variance will be performed with within-subjects factors of stage (iVO_{2max}, $verVO_{2max}$) and sampling time (10 seconds, 30 seconds, and 60 seconds). Post-hoc tests will be performed using pairwise comparisons in SPSS with a Bonferroni correction for main effects. For the interaction effect, paired t-tests with a Bonferroni correction will be performed on the difference between iVO_{2max} and $verVO_{2max}$ for each sampling time. For all three sampling times, sensitivity and specificity of the following criteria for VO_{2max} will be calculated:

plateau (< 150 ml/min increase in VO_2 over the final 2 stages), HR (achievement of at least 90% of age-predicted maximal heart rate), and RER ($\geq$ 1.10). verVO_{2max} will be considered higher if it exceeds iVO_{2max} by more than 2% as suggested by Midgley et al (Midgley, et al., 2007). Sensitivity will be calculated by taking the number of True Positives (Criteria indicates VO_{2max} and iVO_{2max} is within 2% of verVO_{2max}) divided by True Positives plus False Negatives (Criteria not achieved and iVO_{2max} is within 2% of verVO_{2max}). Specificity will be determined by the number of True Negatives (Criteria not achieved and verVO_{2max} greater than 2% higher than iVO_{2max}) divided by the number of True Negatives plus False Positives (Criteria achieved but verVO_{2max} greater than 2% higher than iVO_{2max}).

Chapter III.

Manuscript

Introduction

For decades, cardiorespiratory fitness has been determined through maximal oxygen consumption (VO_{2max}) testing. Despite its widespread use, standard criteria used to verify that true VO_{2max} is reached are lacking. It has been suggested that common primary and secondary criteria are ineffective and vary by too large a margin from person to person to be applied universally[10, 13]. A verification stage completed after a continuous GXT has been shown to elicit similar values to those achieved in a GXT, and thus can act as a way to verify that test subjects have achieved their true VO_{2max}[6].

Several studies have found that when a subject completes a VO_{2max} test, a subsequent verification stage results in subjects either reaching or exceeding the stage in which they originally stopped the test[6, 9, 11]. Furthermore, during verification stages, subjects often attain both HR and RER values similar to or higher than during the original GXT test. Mier and colleagues investigated whether it was necessary to employ a supramaximal verification stage in college athletes who did not achieve a VO_{2max} plateau during a GXT. Researchers found that the VO_{2max} values from the verification stage were not significantly different than the VO_{2max} values from the continuous GXT[11] indicating that similar VO_2 is achieved in the short time period that comprises the verification stage. Similarly, Foster and colleagues had test subjects perform a cycling GXT test, followed by a supramaximal verification stage. This verification stage resulted in the same average VO_{2max} values as the continuous GXT, suggesting true VO_{2max} values can be achieved during a verification stage[6] and that verification stages can be used to verify whether a test subject achieved a maximal value in their test. This same study observed similar findings in runners who performed a treadmill GXT[6]. Furthermore, Midgley et al., have argued the use of a verification stage will increase the probability VO_{2max} is achieved during the

test, as it provides an additional opportunity for the test subject to reach the upper limits of their oxygen uptake. Practically speaking, a verification stage will either verify the attainment of a VO_{2max} during the GXT or will result in a higher measurement for VO_{2max}.

Higher incidences of a plateau during VO_{2max} testing have been observed for 11- and 15-seconds sampling times when compared to 30 second sampling averages for breath by breath VO_2 measurements[2]. Furthermore, 15 and 30s intervals have shown to result in higher VO_{2max} values than 60s interval[1,14]. Finally, one of the issues that could affect the efficacy of a verification stage is that subjects could fatigue quickly due to prior activity and not enough time would be provided for VO_2 to reach VO_{2max}. It could be that the use of shorter sampling times would limit this concern, as the subject would only need to reach VO_{2max} for a shorter window of time. Thus, a greater portion of people would potentially exceed the highest VO_2 achieved during the GXT. Therefore, the need for a verification stage may depend upon the duration of the sampling time.

Despite the widespread use of VO_{2max} testing, there is limited data in the literature assessing the sensitivity and specificity of traditional criteria. Bhammar et al. found poor sensitivity and specificity for traditional VO_2 criteria. However, this study was performed with a limited number of test subjects and the test subjects were children. Thus, these findings are limited in their generalizability[4]. Furthermore, it is not known if the use of different sampling times will impact the effectiveness of traditional primary and secondary criteria for determining VO_{2max} from a continuous GXT. It is important to study the sensitivity and specificity of VO_{2max} criteria as Poole and Jones have suggested that by increasing the use of secondary criteria may lead to an increase in the likelihood of both false negatives and false positives[13]. Additionally, the use of sensitivity and specificity to assess the suitability of primary and secondary criteria for VO_{2max} testing would improve the objective evaluation of these criteria as they measure the degree to which false negatives, as well as false positives, occur.

The purposes of this present study were to: 1) determine the influence of sampling time on the efficacy of a verification stage; and 2) to determine the sensitivity and specificity of primary and secondary VO_{2max} test criteria. It is hypothesized that verification stage VO_{2max} will exceed the VO_{2max} reached in an incremental, graded exercise test in proportion to sampling time and that traditional primary and secondary criteria for achievement of VO_{2max} will exhibit poor sensitivity and specificity regardless of sampling time.

Methods

Subjects: This study evaluated 29 test subjects (14 men, 15 women) with an average age of 21.3 ± 1.2 yr. Before testing, subjects completed a health questionnaire and had no known CV, metabolic, or renal diseases. Additionally, test subjects had no known injuries or other health concerns that would preclude them from exercise, or limit their ability to perform a maximal GXT. Height and weight were measured on the day of testing (men = 179.1 ± 8.5 cm, 80.9 ± 10.4 kg, women = 160.3 ± 5.4 cm, 60.3 ± 5.4 kg) before the VO_{2max} treadmill test.

Treadmill Test: All subjects were monitored for oxygen uptake (VO_2) and respiratory exchange ratio (RER) with a VMax metabolic cart (CareFusion; San Diego, CA) throughout the duration of the test. A Polar heart rate monitor (Lake Success, NY) was utilized to measure heart rate throughout the test. After an initial stage at 3.0 mph and 0% grade, the treadmill speed was increased by 0.5 mph each minute until a speed of 6.0 mph was achieved. After this, the incline of the treadmill was increased by 3% every minute until volitional exhaustion. Subjects were then allowed to walk for 5 minutes at a speed 3.0 mph and a 0% grade. After this rest period, the verification stage was initiated by increasing the speed and grade to the values of the stage preceding the test subject's prior maximal effort. The test then proceeded as described previously until volitional exhaustion.

Statistical Analyses: A repeated measure of analysis of variance was performed with within-subjects' factors of stage (iVO_{2max}, $verVO_{2max}$) and sampling time (10 seconds, 30 seconds, and 60 seconds). Post-hoc tests were performed using pairwise comparisons in SPSS. For the interaction effect, paired t-tests with a Bonferroni correction were performed on the difference between iVO_{2max} and $verVO_{2max}$ for each sampling time. For all three sampling times, sensitivity and specificity of the following criteria for VO_{2max} were calculated: plateau (< 150 ml/min increase in VO_2 over the final 2 stages), HR + RER (achievement of at least 90% of age-predicted maximal heart rate and RER $\geq$ 1.10). $verVO_{2max}$ was considered to be higher if it exceeded iVO_{2max} by more than 2% as suggested by Midgley et al[10]. Sensitivity was calculated by taking the number of True Positives (Criteria indicates VO_{2max} and iVO_{2max} is within 2% of $verVO_{2max}$) divided by True Positives plus False Negatives (Criteria not achieved and iVO_{2max} was within 2% of $verVO_{2max}$). Specificity was determined by the number of True Negatives (Criteria not achieved and $verVO_{2max}$ greater than 2% higher than iVO_{2max}) divided by the number of True Negatives plus False Positives (Criteria achieved but $verVO_{2max}$ greater than 2% higher than iVO_{2max}).

Results

Table 1 shows the effect of sampling time on average values for VO_{2max}. As sampling time increased from 10 seconds to 60 seconds, average VO_{2max} values significantly ($P < 0.05$) decreased (10s > 30s > 60s). The difference between $verVO_{2max}$ and iVO_{2max} was greater for the 10 second sampling time than for the 60 second sampling time.

Table 2 shows the sensitivity and specificity for VO_{2max} plateau. Sensitivity for the incidence of a plateau for VO_{2max} was $\leq$ 30% for 10-, 30-, and 60-second sampling times. Specificity was < 45% for all three sampling times.

Table 3 demonstrates the sensitivity and specificity of HR + RER secondary criteria. Sensitivity for HR + RER was > 84% for 10-, 30-, and 60-second sampling times. The highest sensitivity was observed with the 30-second sampling time. For all three sampling times, specificity was < 28%.

Discussion

In this study, we observed shorter sampling times resulted in higher VO_{2max} values when compared to longer sampling times. In support of this, Astorino et al. found that as the sampling time decreased, both VO_{2max} and the incidence of a plateau increased[1]. Our data also suggest there is a greater need for verification stages when shorter sampling times are implemented. The duration of our verification stage was typically one to two minutes. Because VO_2 is increasing at the onset of a verification stage, it is possible the 60s sampling time includes several data points that do not reflect the steady-state VO_2 associated with the work rate and thus fails to deliver an average VO_2 that truly reflects VO_{2max}. This concern becomes even greater if longer sampling times are used. However, it should also be realized that sampling variability has a greater impact with shorter sampling times. For example, an aberrant data point would have a larger influence on the average of 10 breaths than the average of 60 breaths.

Similar to the present study, several studies have found verification stages yield VO_{2max} values comparable to those achieved during a continuous GXT. Foster et al. observed similar values for VO_{2max} in a verification stage and during GXT's in which a plateau was evident. This was true for both treadmill and cycling tests[6]. Midgley et al. found no statistically significant differences between VO_{2max} values during a running GXT and a verification stage[9]. Therefore, it appears verification stages are effective at confirming VO_2 values achieved during a GXT. In the current study, the verification stage resulted in a higher VO_{2max} in 31-62% of our tests (depending on the sampling time), which is perhaps more meaningful than the observation of similar aggregate values for iVO_{2max} and $verVO_{2max}$.

Practically speaking, this represents a relatively substantial portion of VO_{2max} tests and illustrates the utility of verification stages beyond the need to simply confirm the value from the GXT.

We observed poor sensitivity and specificity for the use of a VO_2 plateau for confirming VO_{2max} was achieved during the GXT. This finding was not impacted substantially with sampling duration. We expected there would not be a high degree of sensitivity for a plateau as it occurs only 15% of the time in non-athletes and ~50% of the time in athletes[6]. Howley et al. proposed these numbers may be even lower as children, sedentary, and elderly populations have a harder time achieving a plateau, especially as cut-offs for VO_2 changes to indicate a plateau compose a large portion of their VO_{2max}[8]. However, the finding of low specificity was surprising as it suggests a high number of test subjects that exhibit a plateau do not actually achieve a VO_{2max} during the incremental GXT. Few studies have evaluated the sensitivity and specificity of a plateau. Bhammar and colleagues found similar values for sensitivity and specificity for plateau as the current study in nonobese and obese children (22 and 44% respectively). Glassford et al. compared VO_{2max} values with various cut-off values of 150 mL/min[16], 54 mL/min[3], and 80 mL/min[12]. They found no difference in VO_{2max} values between the treadmill tests. However, even though criteria for the plateaus was met, some subjects still experienced a significant increase in VO_{2max} when given a higher work rate[7, 8]. Therefore, the incidence of a plateau, as traditionally employed, may not be a good criterion to validate VO_{2max} measurements.

Our data showed a high sensitivity and low specificity for the use of combined HR + RER criteria, which did not appear to substantially change with sampling time. This suggests reaching these criteria is a common occurrence. However, these secondary criteria were ineffective at parsing out those whose who did not achieve VO_{2max}. Multiple studies have described flaws with secondary criteria. Howley et al. found that not all test subjects are able to achieve RER criteria, especially in young and elderly populations[8]. They also suggested the large standard deviations associated with

estimated maximal HR_{max} values may limit the success of heart rate values as secondary criteria[8]. Poole & Jones stated age-predicted HR_{max} has a confidence interval of ± 35 bpm and also reported that it had been recorded as high as ± 45 bpm in some studies[13]. Cumming and Borysyk also concluded that the heart rate range for a maximal heart rate in test subjects is too wide and too large of an inter-subject variability exists for HR values to be used as a criterion for VO_{2max}[5]. Therefore, the secondary criteria as employed in the current study also appear to be ineffective for validating VO_{2max}.

This study had a few limitations, such as a relatively small sample size. In particular, a larger sample size might have provided more precision for the sensitivity and specificity measurements. Post-hoc power calculations suggest having a higher number of test subjects could contribute to detecting a main effect for stage, as the power was 0.10. Additionally, it is possible that the use of a plateau or the use of HR + RER values (secondary criteria) are good for identifying VO_{2max} but we failed to identity the right cut-off values for these criteria. However, the specific primary and secondary criteria were chosen because they are commonly used in research. Midgley et al. identified that the most common secondary criteria include achieving a HR value of 95% age-predicted heart rate and an RER value of equal to or > 1.15[10]. Furthermore, it was observed that the plateau criteria used in the present study is the one most commonly utilized in the literature[10]. Thus, our findings show common criteria that would be considered acceptable in the literature have poor sensitivity and/or specificity. The characteristics of an optimal verification stage are currently unknown. However, the protocol employed in the current study with respect to both the recovery time and the intensity is within the suggested ranges[15]. Improvements in the verification stage protocol would likely lead to an increased verVO_{2max} and would result in an even greater proportion of tests that required a verification stage and even worse sensitivity/specificity for traditional criteria.

Manuscript References

1. Astorino TA. Alterations in VO2max and the VO2 Plateau With Manipulation of Sampling Interval. *Clinical Physiology and Functional Imaging*. 2008;(29):60-67. doi:Clin Physiol Funct Imaging (2009) 29, pp60–67 doi: 10.1111/j.1475-097X.2008.00835.x.
2. Astorino, T.A., Robergs, R.A., Ghiasvand, F., Marks, D., & Burns, S. (2000) Incidence of the oxygen plateau at VO2max during exercise testing to volitional fatigue. *Journal of Exercise Physiology Online*, *3*(4), pp. 1-12.
3. Astrand I. Aerobic work capacity in men and women with special reference to age. *Acta Physiologica-The Scandinavian Physiological Society*. 1960;49(169):1-92.
4. Bhammar DM, Stickford JL, Bernhardt V, Babb TG. Verification of Maximal Oxygen Uptake in Obese and Nonobese Children. *Medicine & Science in Sports & Exercise*. 2017;49(4):702-710. doi:10.1249/mss.0000000000001170.
5. Cumming GR, Borysyk LM. Criteria for maximum oxygen uptake in men over 40 in a population survey. *Medicine & Science in Sports & Exercise*. 1972;4(1):18-22.
6. Foster, C., Kuffel, E., Bradley, N., Battista, R. A., Wright, G., Porcari, J. P., . . . DeKoning, J. J. (2007). VO2max during successive maximal efforts. European Journal of Applied Physiology, 102, 67-72.
7. Glassford RG, Baycroft GHY, Sedgwick AW, Macnab RBJ. Comparison of maximal oxygen uptake values determined by predicted and actual methods. *Journal of Applied Physiology* . May 1965. doi:https://doi.org/10.1152/jappl.1965.20.3.509.
8. Howley ET, Bassett DR, Welch HG. Criteria for maximal oxygen uptake. Medicine & Science in Sports & Exercise. 1995;27(9):1292-1299. doi:10.1249/00005768-199509000-00009.
9. Midgley AW, Mcnaughton LR, Carroll S. Verification phase as a useful tool in the determination of the maximal oxygen uptake of distance runners. Applied Physiology, Nutrition, and Metabolism. 2006;31(5):541-548. doi:10.1139/h06-023.
10. Midgley, A. W., Mcnaughton, L. R., Polman, R., & Marchant, D. (2007). Criteria for Determination of Maximal Oxygen Uptake. Sports Medicine, 37(12), 1019-1028. doi:10.2165/00007256-200737120-00002
11. Mier, C. M., Alexander, R. P., & Mageean, A. L. (2012). Achievement of VO2Max Criteria During a Continuous Graded Exercise Test and a Verification Stage Performed by College Athletes. Journal of Strength and Conditioning Research,26(10), 2648-2654. doi:10.1519/jsc.0b013e31823f8de
12. Mitchell JH, Sproule BJ, Chapman CD. The physiological meaning of the maximal oxygen intake test. *Journal of Clinical Investigation*. 1958;37(4):538-547.
13. Poole, D. C., & Jones, A. M. (2017). Measurement of the maximum oxygen uptake V̇o2max: V̇o2peak is no longer acceptable. Journal of Applied Physiology,122(4), 997-1002. doi:10.1152/japplphysiol.01063.2016
14. Robergs RA. AN EXERCISE PHYSIOLOGIST'S "CONTEMPORARY" INTERPRETATIONS OF THE "UGLY AND CREAKING EDIFICES" OF THE VO2MAX CONCEPT. *Official Journal of The American Society of Exercise Physiologists* . 2001;4(1).
15. Schaun GZ. The Maximal Oxygen Uptake Verification Phase: a Light at the End of the Tunnel? *Sports Medicine - Open*. 2017;3(1). doi:10.1186/s40798-017-0112-1.

16. Taylor HL, Buskirk E, Henschel A. Maximal Oxygen Intake as an Objective Measure of Cardio-Respiratory Performance. *Journal of Applied Physiology*. 1955;8(1):73-80. doi:10.1152/jappl.1955.8.1.73.

Tables and Figures

Table 1. Average (± SD) iVO_{2max} and $verVO_{2max}$ along with the percentage of tests in which $verVO_{2max} > iVO_{2max}$. *-Main effect for sampling time (10s > 30s > 60s, $P < 0.05$), †-Significant stage x sample time interaction ($verVO_{2max} – iVO_{2max}$ for 10s > 60s, $P < 0.05$)

Sampling Time*	iVO_{2max} (ml/kg/min)	$verVO_{2max}$ (ml/kg/min)	$verVO_{2max} > iVO_{2max}$ (%)†
10s	47.9 ± 8.31	48.85 ± 7.97	62%
30s	46.94 ± 8.62	47.28 ± 7.97	41%
60s	46.17 ± 8.62	46.00 ± 8.00	31%

Table 2. Sensitivity and specificity of VO_{2max} primary criteria (plateau).

	10s	30s	60s
Sensitivity	18.2	23.5	30
Specificity	38.9	33.3	44.4

Table 3. Sensitivity and specificity of secondary criteria (HR + RER).

	10s	30s	60s
Sensitivity	90.9	100	84.2
Specificity	17.7	27.3	11.1

Appendix A

Informed Consent Form

Consent to Participate in Research

Identification of Investigators & Purpose of Study

You are being asked to participate in a research study conducted by Emily Kontos and Chris Womack from James Madison University. The purpose of this study is to determine is to determine whether a post-test verification stage will improve the determination of maximal oxygen consumption. This study will contribute to completion of Emily Kontos' Honors Thesis.

Research Procedures

Should you decide to participate in this research study, you will be asked to sign this consent form once all your questions have been answered to your satisfaction. In addition, you will be asked to complete a health history questionnaire that includes your known history of disease, medical procedures, and medications. This study consists of one visit to the Human Performance Laboratory in Godwin Hall, Room 209. You will perform a treadmill test that allows us to determine the maximal amount of oxygen that your body is capable of using (VO2max). This is a good evaluation of your cardiovascular system's ability to supply blood to your working muscles. For this visit, you will be asked to refrain from eating or drinking anything except water for three hours prior to the test.

During the test, you will begin walking on a treadmill at 3.0 miles/hour. The speed of the treadmill will increase every minute until you reach 6.0 miles/hour. After that, the elevation (grade) of the treadmill will increase by 3% per minute until you indicate that you can no longer continue. After a 5-minute rest, you will resume the test at the intensity that preceded your highest intensity achieved during the test. We will continue the test in the same manner until you indicate that you can no longer continue. Throughout the test, you will be breathing through a mouthpiece so that we can collect and analyze your expired air for oxygen content. You will also wear a strap around your chest so that we can monitor your heart rate.

Time Required

Your participation will require one session that will take about 45 minutes.

Risks

Research on risk of exercise testing has suggested that approximately six cardiac events occur for every 10,000 exercise tests. The risk of death is even less, with a rate of approximately one death per 1,000,000 tests. This is likely to be even less in college-aged individuals. In the unlikely event of an event, at least one investigator will be CPR-trained at every test.

Benefits

Potential benefits from participation in this study include feedback on your current level of cardiorespiratory fitness. In addition to your actual scores, you will be given established norms for both fitness-related variables.

Confidentiality

The results of this research will be presented at relevant regional and national/international conferences. Our findings will also be published in relevant research journals and/or books in the field of exercise science. The results of this project will be coded in such a way that your identity will not be attached to the final form of this study. The researcher retains the right to use and publish non-identifiable data. While individual responses are confidential, aggregate data will be presented representing averages or generalizations about the responses as a whole. All data will be stored in asecure location accessible only to the researcher. Upon completion of the study, all information that matches up individual respondents with their answers will be destroyed.

Participation & Withdrawal

Your participation is entirely voluntary. You are free to choose not to participate. Should you choose to participate, you can withdraw at any time without consequences of any kind.

Questions about the Study

If you have questions or concerns during the time of your participation in this study, or after its completion or you would like to receive a copy of the final aggregate results of this study, please contact:

Emily Kontos

Department of Kinesiology

James Madison University

kontosej@dukes.jmu.edu

Christopher Womack

Department of Kinesiology

James Madison University

womackcx@jmu.edu

Telephone: (540) 568-6515

Questions about Your Rights as a Research Subject

Dr. Taimi Castle

Chair, Institutional Review Board

James Madison University

(540) 568-5929 castletl@jmu.edu

Giving of Consent

I have read this consent form and I understand what is being requested of me as a participant in this study. I freely consent to participate. I have been given satisfactory answers to my questions. The investigator provided me with a copy of this form. I certify that I am at least 18 years of age.

_____

Name of Participant (Printed)

_____ _____

Name of Participant (Signed) Date

_____ __

Appendix B

VO_{2max} Data Sheet

Subject #- Temp-

Date- Rh-

Height (cm)- Pb-

Weight (kg)- Region %Fat- BMD t-score-

Age-

Time	Speed (mph)	Elevation (%)	HR	RPE	VO_2 (L/min)
1	3.0	0			
2	3.5	0			
3	4.0	0			
4	4.5	0			
5	5.0	0			
6	5.5	0			
7	6.0	0			
8	6.0	3			
9	6.0	6			
10	6.0	9			
11	6.0	12			
12	6.0	15			
13	6.0	18			
14	6.0	21			
15	6.0	24			

Exercise test duration:

Difference between VO_2 in last 2 full stages:

Validation Stage

Time	Speed	Elevation	VO_2

Validation Stage duration:

VO_{2max} from validation stage? (Y/N):

MAX HR	MAX VO_2 (L/min)	MAX RQ	MAX RPE

References

1. Astorino TA. Alterations in VO2max and the VO2 Plateau With Manipulation of Sampling Interval. *Clinical Physiology and Functional Imaging*. 2008;(29):60-67. doi:Clin Physiol Funct Imaging (2009) 29, pp60–67 doi: 10.1111/j.1475-097X.2008.00835.x.
2. Astorino, T.A., Robergs, R.A., Ghiasvand, F., Marks, D., & Burns, S. (2000) Incidence of the oxygen plateau at VO2max during exercise testing to volitional fatigue. *Journal of Exercise Physiology Online*, *3*(4), pp. 1-12.
3. Astrand I. Aerobic work capacity in men and women with special reference to age. *Acta Physiologica-The Scandinavian Physiological Society*. 1960;49(169):1-92.
4. Bhammar DM, Stickford JL, Bernhardt V, Babb TG. Verification of Maximal Oxygen Uptake in Obese and Nonobese Children. *Medicine & Science in Sports & Exercise*. 2017;49(4):702-710. doi:10.1249/mss.0000000000001170.
5. Cumming GR, Borysyk LM. Criteria for maximum oxygen uptake in men over 40 in a population survey. *Medicine & Science in Sports & Exercise*. 1972;4(1):18-22.
6. Duncan GE, Howley ET, Johnson BN. Applicability of ??VO2max criteria: discontinuous versus continuous protocols. Medicine & Science in Sports & Exercise. 1997;29(2):273-278. doi:10.1097/00005768-199702000-00017.
7. Foster, C., Kuffel, E., Bradley, N., Battista, R. A., Wright, G., Porcari, J. P., . . . DeKoning, J. J. (2007). VO2max during successive maximal efforts. European Journal of Applied Physiology, 102, 67-72.
8. Glassford RG, Baycroft GHY, Sedgwick AW, Macnab RBJ. Comparison of maximal oxygen uptake values determined by predicted and actual methods. *Journal of Applied Physiology* . May 1965. doi:https://doi.org/10.1152/jappl.1965.20.3.509.
9. HILL, A. V., and H. LUPTON. Muscular exercise, lactic acid, and the supply and utilization of oxygen. Q. J. Med. 16:135–171, 1923.
10. Howley ET, Bassett DR, Welch HG. Criteria for maximal oxygen uptake. Medicine & Science in Sports & Exercise. 1995;27(9):1292-1299. doi:10.1249/00005768-199509000-00009.
11. Juranovich, G. D. (2001). The specificity and sensitivity of the criteria used to measure VO2Max. https://ro.ecu.edu.au/theses_hons/555
12. Midgley AW, McNaughton LR, Carroll S. Time at VO_{2max} during intermittent treadmill running: test protocol dependent or methodological artefact? Int J Sports Med. Epub 2007 May 11
13. Midgley AW, Mcnaughton LR, Carroll S. Verification phase as a useful tool in the determination of the maximal oxygen uptake of distance runners. Applied Physiology, Nutrition, and Metabolism. 2006;31(5):541-548. doi:10.1139/h06-023.
17. Midgley, A. W., Mcnaughton, L. R., Polman, R., & Marchant, D. (2007). Criteria for Determination of Maximal Oxygen Uptake. Sports Medicine, 37(12), 1019-1028. doi:10.2165/00007256-200737120-00002
18. Mier, C. M., Alexander, R. P., & Mageean, A. L. (2012). Achievement of VO2Max Criteria During a Continuous Graded Exercise Test and a Verification Stage Performed by College Athletes. Journal of Strength and Conditioning Research,26(10), 2648-2654. doi:10.1519/jsc.0b013e31823f8de

19. Mitchell JH, Sproule BJ, Chapman CD. The physiological meaning of the maximal oxygen intake test. *Journal of Clinical Investigation*. 1958;37(4):538-547.
20. Nes, B. M., Vatten, L. J., Nauman, J., Janszky, I., & Wisløff, U. (2014). A Simple Nonexercise Model of Cardiorespiratory Fitness Predicts Long-Term Mortality. Medicine & Science in Sports & Exercise, 46(6), 1159-1165. doi:10.1249/mss.0000000000000219
21. Noakes TD, StClair Gibson A (2004) Logical limitation in the "catastrophe" models of fatigue during exercise in humans. Br J Sports Med 38:648–649
22. Poole, D. C., & Jones, A. M. (2017). Measurement of the maximum oxygen uptake V̇o2max: V̇o2peak is no longer acceptable. Journal of Applied Physiology,122(4), 997-1002. doi:10.1152/japplphysiol.01063.2016
23. Robergs RA. AN EXERCISE PHYSIOLOGIST'S "CONTEMPORARY" INTERPRETATIONS OF THE "UGLY AND CREAKING EDIFICES" OF THE VO2MAX CONCEPT. *Official Journal of The American Society of Exercise Physiologists* . 2001;4(1).
24. Schaun GZ. The Maximal Oxygen Uptake Verification Phase: a Light at the End of the Tunnel? *Sports Medicine - Open*. 2017;3(1). doi:10.1186/s40798-017-0112-1.
25. Taylor HL, Buskirk E, Henschel A. Maximal Oxygen Intake as an Objective Measure of Cardio-Respiratory Performance. *Journal of Applied Physiology*. 1955;8(1):73-80. doi:10.1152/jappl.1955.8.1.73.
26. Wei M. Relationship Between Low Cardiorespiratory Fitness and Mortality in Normal-Weight, Overweight, and Obese Men. Jama 1999;282(16):1547
27. Yoon, B., Kravitz, L., & Robergs, R. (2007). VO2max, Protocol Duration, and the VO2 Plateau. Medicine & Science in Sports & Exercise, 39(7), 1186-1192. doi:10.1249/mss.0b13e318054e304