OPEN

Verification of Maximal Oxygen Uptake in Active Military Personnel During Treadmill Running

Peter S. Figueiredo,[1,2] David P. Looney,[1] J. Luke Pryor,[3] Elizabeth M. Doughty,[1,2] Holly L. McClung,[1] Sai V. Vangala,[1,2] William R. Santee,[1,2] Beth A. Beidleman,[1] and Adam W. Potter[1]

[1]Biophysics and Biomedical Modeling Division, U.S. Army Research Institute of Environmental Medicine (USARIEM), Natick, Massachusetts; [2]Oak Ridge Institute for Science and Education (ORISE), Oak Ridge, Tennessee; and [3]Center for Research and Education in Special Environments, Department of Exercise and Nutrition Sciences, University at Buffalo, Buffalo, New York

Abstract

Figueiredo, PS, Looney, DP, Pryor, JL, Doughty, EM, McClung, HL, Vangala, SV, Santee, WR, Beidleman, BA, and Potter, AW. Verification of maximal oxygen uptake in active military personnel during treadmill running. *J Strength Cond Res* 36(4): 1053–1058, 2022—It is unclear whether verification tests are required to confirm "true" maximal oxygen uptake ($\dot{V}O_2$max) in modern warfighter populations. Our study investigated the prevalence of $\dot{V}O_2$max attainment in U.S. Army soldiers performing a traditional incremental running test. In addition, we examined the utility of supramaximal verification testing as well as repeated trials for familiarization for accurate $\dot{V}O_2$max assessment. Sixteen U.S. Army soldiers (1 woman, 15 men; age, 21 ± 2 years; height, 1.73 ± 0.06 m; body mass, 71.6 ± 10.1 kg) completed 2 laboratory visits, each with an incremental running test (modified Astrand protocol) and a verification test (110% maximal incremental test speed) on a motorized treadmill. We evaluated $\dot{V}O_2$max attainment during incremental testing by testing for the definitive $\dot{V}O_2$ plateau using a linear least-squares regression approach. Peak oxygen uptake ($\dot{V}O_2$peak) was considered statistically equivalent between tests if the 90% confidence interval around the mean difference was within ±2.1 ml·kg^{-1}·min^{-1}. Oxygen uptake plateaus were identified in 14 of 16 volunteers for visit 1 (87.5%) and all 16 volunteers for visit 2 (100%). Peak oxygen uptake was not statistically equivalent, apparent from the mean difference in $\dot{V}O_2$peak measures between the incremental test and verification test on visit 1 (2.3 ml·kg^{-1}·min^{-1}, [1.3–3.2]) or visit 2 (1.1 ml·kg^{-1}·min^{-1} [0.2–2.1]). Interestingly, $\dot{V}O_2$peak was equivalent, apparent from the mean difference in $\dot{V}O_2$peak measures between visits for the incremental tests (0.0 ml·kg^{-1}·min^{-1} [−0.8 to 0.9]) but not the verification tests (−1.2 ml·kg^{-1}·min^{-1} [−2.2 to −0.2]). Modern U.S. Army soldiers can attain $\dot{V}O_2$max by performing a modified Astrand treadmill running test. Additional familiarization and verification tests for confirming $\dot{V}O_2$max in healthy active military personnel may be unnecessary.

Key Words: verification phase, familiarization, cardiopulmonary exercise test, plateau

Introduction

Dating back to the early work of Hill and Lupton (14), maximal oxygen uptake ($\dot{V}O_2$max) has remained the gold standard for quantifying maximal O_2 diffusion, transport, and utilization by the body and a paramount measure for both health (26,35) and aerobic fitness (5). Testing for $\dot{V}O_2$max generally requires an incremental exercise protocol lasting 8–12 minutes that starts at moderate intensity and continues until exhaustion (8). The unequivocal criterion for $\dot{V}O_2$max attainment is a clearly displayed plateau of $\dot{V}O_2$ in relation to increasing work rate (14,38). If an individual ends testing before displaying a clear $\dot{V}O_2$ plateau, the highest achieved $\dot{V}O_2$ is instead deemed $\dot{V}O_2$peak (23). Many individuals do not achieve the $\dot{V}O_2$ plateau before exhaustion (12); therefore, secondary criteria for supporting maximal effort during testing have been developed (25). However, investigators have scrutinized these secondary criteria because of the large individual variability in these physiological markers at maximal exercise (23,25,31).

In view of this, Poole and Jones (31) advocated for inclusion of an additional verification test at ~110% of the highest work rate performed during the initial test to confirm $\dot{V}O_2$max attainment. This verification test is intended to determine whether the subject can exceed the highest $\dot{V}O_2$ achieved on the initial incremental test with another work rate increase. However, Green and Askew (12) argued that low $\dot{V}O_2$ plateau incidence rates and unreliable $\dot{V}O_2$peak data are often caused by lack of motivation and/or test naivety. These authors instead advocated for further emphasis on familiarization procedures.

Previous studies on $\dot{V}O_2$max verification strategies have shown utility in different populations (12,39), with a major focus on trying to elucidate the minimal changes in $\dot{V}O_2$max of highly trained athletes. There have been no investigations on warfighters to date. Long-term declines in aerobic exercise performance in U.S. Army soldiers have been observed over the past few decades (17–19). Although longitudinal studies (18,19) did not determine a concurrent decrease in $\dot{V}O_2$max of soldiers, only 3 group mean values from studies from 1975 to 1998 were analyzed. Recent studies have emphasized the influence of motivation on physical fitness test performance in warfighter populations (7,19,40). Notably, Buch et al. (7) identified intrinsic motivation as the strongest predictor of longitudinal $\dot{V}O_2$max improvement in Norwegian military cadets. Although warfighters often serve as human research volunteers, they are less likely to be familiar with $\dot{V}O_2$max

testing procedures compared with athletes. As such, it is unclear whether modern warfighters require additional verification and familiarization tests to attain a "true" $\dot{V}O_2$max.

Our study addressed 3 specific research aims regarding strategies for confirming $\dot{V}O_2$max attainment in active military personnel. Primarily, we investigated the prevalence of the definitive $\dot{V}O_2$ plateau in U.S. Army soldiers when performing a traditional incremental running test. Subsequently, we examined whether verification testing at supramaximal intensity could elicit $\dot{V}O_2$peak results comparable to the traditional incremental test. Finally, we assessed the effect of familiarization by repeating these tests on a second visit and comparing $\dot{V}O_2$peak between visits. This study evaluated the necessity of additional verification and familiarization procedures for accurate determination of $\dot{V}O_2$max in modern warfighters.

Methods

Experimental Approach to the Problem

We conducted a within-subject design study to evaluate strategies for determining $\dot{V}O_2$max attainment in active military personnel. Study volunteers were required to complete 2 laboratory visits (visits 1 and 2) each with an incremental treadmill running test followed by a supramaximal verification test. We used a $\dot{V}O_2$ plateau identification method, specific to the individual's $\dot{V}O_2$-workrate slope, to identify incidences of $\dot{V}O_2$ plateau during incremental running tests. We evaluated the utility of verification testing by comparing achieved $\dot{V}O_2$peak with those attained during the preceding incremental tests. Comparisons in $\dot{V}O_2$peak between incremental tests on visit 1 and visit 2 were analyzed to assess the effect of familiarization.

Subjects

Sixteen U.S. Army human research volunteers (18–25; 1 woman, 15 men; age, 21 ± 2 years; height, 1.73 ± 0.06 m; body mass, 71.6 ± 10.1 kg; Mean ± SD) naive to $\dot{V}O_2$max testing participated in this study. Subjects' most recent Army Physical Fitness Test (APFT) score was above the 70th percentile for male soldiers between 17 and 21 years old (push-up, 58 ± 13 reps; sit-up; 64 ± 10 reps; 2-mile run, 13:56 ± 1:16) (16). Specific to cardiorespiratory fitness, volunteers had a higher $\dot{V}O_2$max (49.4 ± 4.5 ml·kg⁻¹·min⁻¹) than the 50th percentile (48 ml·kg⁻¹·min⁻¹) listed for 20- to 29-year-old men in the Fitness Registry and the Importance of Exercise National Database (15). Each volunteer was briefed on the purpose, risks, and benefits of the study before providing written informed consent. To be eligible for this study, subjects were required to be between 18 and 44 years old, exercise for at least 30 minutes on at least 2 days per week, weigh less than 128 kg, and were free of any musculoskeletal injuries, illnesses, or medical conditions that compromise the ability to exercise. In addition, subjects who had difficulty breathing into a mouthpiece or claustrophobia, any history of gastrointestinal disease or surgery, or were pregnant were excluded from participation. This study was approved by the institute's scientific review committee and by the institutional review board at the U.S. Army Medical Research and Development Command (USAMRDC; Ft. Detrick, MD).

Procedures

Each volunteer attended 2 morning laboratory visits scheduled at the same time of day with at least 2 recovery days in between. Before each visit, volunteers were instructed to avoid high-intensity exercise, including resistance training (>48 hours), alcohol (>24 hours), as well as caffeine, nicotine, and food (>10 hours). Outside of study restrictions, volunteers were asked to maintain their normal dietary habits. To ensure proper hydration level, volunteers were provided bottled water (500 ml) to drink the night before and the morning of each visit. Adequate hydration was confirmed by checking that a urine sample provided by the subject had a specific gravity ≤1.030. Volunteers wore standard physical training attire (shorts, t-shirt, socks, and running shoes).

After measures of height and body mass, volunteers were fitted with a chest belt physiological status monitor system (EQ02; Hidalgo Ltd., Cambridge, United Kingdom) that measured heart rate (HR) and donned a respirometer mask connected to an open circuit spirometry unit (ParvoMedics TrueOne 2,400; Parvo-Medics; Salt Lake City, UT). The laboratory open circuit spirometry unit was warmed up for >60 minutes with at least 2 flowmeter and gas analyzer calibrations before testing in accordance with manufacturer instructions. Volunteers warmed up with an incremental treadmill walk before performing a modified Astrand running protocol (29) for the incremental running test. Each incremental treadmill walk began with a 3-minute stage at 1.16 m·s⁻¹ on a 0% incline. The treadmill speed was increased by 0.09 m·s⁻¹ every 2 minutes thereafter until reaching the highest speed 1.97 m·s⁻¹.

Each volunteer began the incremental running tests by running 3 minutes at a speed based on their self-reported 2-mile run pace (2.74 ± 0.19 m·s⁻¹) and the ACSM running equation (1). All volunteers had performed the 2-mile run as part of the APFT within the previous 2 ± 1 months. First, the net oxygen cost of running at the individual's 2-mile run pace was calculated. Then, the incremental treadmill test speed was calculated as the speed that would result in an equivalent net oxygen cost when running on a 10% incline. Volunteers were provided 12-minute rest between tests to better ensure reproducibility of maximal $\dot{V}O_2$ responses (37). During this time, each volunteer unmasked and was free to drink water, sit, stand, or move around leisurely. Heart rate was monitored continuously throughout the rest interval with the lowest value recorded as the recovery HR. Subsequently, the volunteers completed the verification test to volitional exhaustion with the treadmill set to the average incline over the final 2 minutes of the incremental test and the speed increased by 10%. Each volunteer performed all tests under the supervision of the same researcher and were given moderate encouragement (i.e., no yelling or screaming). Physiologic data ($\dot{V}O_2$, respiratory exchange ratio [RER], and HR) were averaged over 30-second epochs. Heart rate was scaled as a percentage of the age-predicted maximal heart rate (APMHR) calculated using the equation from Fox and Haskell (10).

Statistical Analyses

Data were analyzed using R (Version 3.3.1; R Foundation for Statistical Computing; Vienna, Austria) (32) and are displayed as mean and 90% confidence intervals (90% CIs) unless stated otherwise. We used the linear regression approach outlined by Midgley et al. (24) for $\dot{V}O_2$ plateau identification. A simple linear regression model was fit to the $\dot{V}O_2$ data collected from minute −6 to minute −2 before termination of each incremental test. The expected $\dot{V}O_2$peak was considered to be the model's predicted value at minute 0. A $\dot{V}O_2$ plateau was defined as a difference between the expected and measured $\dot{V}O_2$peak that was greater than half the regression slope. Statistical equivalency in $\dot{V}O_2$max

was evaluated between tests using the CI approach for multiple group comparisons (34). Planned contrasts were made for within-visit comparisons (incremental vs. verification) as well as between-visit comparisons for incremental and verification tests, respectively. For these planned contrasts, $\dot{V}o_2$peak was statistically equivalent between tests if the 90% CI was within $\pm$ 2.1 ml kg^{-1} min^{-1} (38). This is equal to an absolute $\dot{V}o_2$peak cutoff of 150.36 ml·min^{-1} when calculated for the average body mass of the current study (71.6 kg). This criterion was selected from Taylor et al. (38), who used a similar, discontinuous protocol. Mean differences and 90% CI for each planned contrast were determined using mixed effect models with random intercepts. A minimum of 13 subjects are necessary to detect statistical equivalence between $\dot{V}o_2$max tests using the CI approach for multiple group comparisons (20) (α, 0.05; β, 0.2; equivalence limit, 2.1 ml·kg^{-1}·min^{-1}) based on previous data (22). Between-visit reliability for the incremental and verification tests was evaluated based on intraclass correlation coefficient (ICC) values (27). Bland-Altman plots of agreement were generated to observe agreement between planned contrasts.

Results

Table 1 presents mean test duration, HRmax, RER, and $\dot{V}o_2$max of volunteers in the incremental running tests and subsequent verification tests at both visit 1 and visit 2. For visit 1 and visit 2, recovery HR was 55% APMHR [52–58] and 51% APMHR [49–54], respectively. Oxygen uptake plateaus were identified in 14 of 16 volunteers for visit 1 (87.5%) and all 16 volunteers for visit 2 (100%). Figure 1 displays Bland-Altman plots of agreement for the planned contrasts. Peak oxygen uptake was not statistically equivalent, as apparent from mean $\dot{V}o_2$peak difference between incremental tests and verification tests on visit 1 (-2.3 ml·kg^{-1}·min^{-1}, [-1.3 to -3.2]) or visit 2 (-1.1 ml·kg^{-1}·min^{-1} [-0.2 to -2.1]). Between-visit comparisons revealed that $\dot{V}o_2$peak was statistically equivalent, as apparent from mean $\dot{V}o_2$peak difference between the incremental tests on visit 1 and visit 2 (0.0 ml·kg^{-1}·min^{-1} [-0.9 to 0.8]) and highly reliable (ICC, 0.95). Conversely, $\dot{V}o_2$peak was not statistically equivalent between verification tests on visit 1 and visit 2 with a mean $\dot{V}o_2$peak difference of (1.2 ml·kg^{-1}·min^{-1} [0.2–2.2]) with lower between-visit reliability (ICC, 0.76).

Discussion

We identified $\dot{V}o_2$ plateau definitive of $\dot{V}o_2$max achievement in almost all incremental exercise tests analyzed in visit 1 and visit 2 (93.75%) using individualized linear regressions described by Midgley et al. (24). This suggests that most soldiers can attain $\dot{V}o_2$max by performing a traditional incremental running test. Despite the work rate increase, verification testing resulted in lower-than-anticipated $\dot{V}o_2$peak. Consequently, these findings

may indicate that extra verification and familiarization strategies are only necessary in circumstances requiring an especially low margin of error.

The modified Astrand protocol used in the current study elicited a $\dot{V}o_2$ plateau in 87.5% of the volunteers on their first visit. Pollock et al. (29) observed similar $\dot{V}o_2$ plateau attainment when comparing the modified Astrand protocol (80%) to the Balke (69%), Bruce (69%), and Ellestad (59%) protocols. Mean $\pm$ SD for the duration of the first incremental test was 580 $\pm$ 48 seconds. Pollock et al. (29) found that within the trained group, the modified Astrand protocol had a similar mean test duration to our study as well as the smallest SD in test duration among the 4 protocols studied (498 $\pm$ 60 seconds). The modified Astrand protocol with a subject-specific starting intensity proves to be reliable in $\dot{V}o_2$ plateau attainment and test duration in trained (29) and active duty military populations. Test durations in the current study and Pollock et al. (29) fall within the 7- to 10-minute recommendation by Astorino et al. (4) in their reinvestigation of optimal $\dot{V}o_2$max test durations. Durations of 7–10 minutes during incremental treadmill running provide the highest $\dot{V}o_2$ and corresponding physiological measures compared with longer tests (4).

Identification of the plateau is prominently impacted by data processing (2,4) and criteria used (24,25,31). Plateau identification using the Taylor et al. (38) criteria is affected by different sampling intervals: 15-second (91%) and 30-second (89%) sampling intervals provide higher incidence of plateau over breath-by-breath (81%) and 60 seconds (59%) (2). Despite 30-second sampling rate, Midgley et al. (24) found that the large between-subject variation around the mean $\dot{V}o_2$-workrate slope renders the satisfaction of the Taylor et al. (38) criteria of 150 ml·min^{-1}, an artifact of differences in individual $\dot{V}o_2$-workrate slopes. The threshold of 150 ml·min^{-1} was between 36 and 90% of the slope for each of their volunteers, which represents small deviations from linearity for some and a plateau for others (24). For this reason, a criterion relative to the individual's own $\dot{V}o_2$-workrate response is more effective at identifying the $\dot{V}o_2$ plateau (24). The Midgley et al. approach (24) also takes 4 minutes of the $\dot{V}o_2$-workrate slope into account which significantly reduces the chances of error from comparing just 2 $\dot{V}o_2$ measures. Researchers can expect that within 1 laboratory visit, active duty military personnel undergoing the modified Astrand protocol will have high incidence of $\dot{V}o_2$ plateau, identified by the Midgley et al. approach (24), despite suspected decreases in motivation within the cohort (7,19,40).

Verification testing within the current study provided lower-than-anticipated $\dot{V}o_2$peak measures. Possible contributors to the lower $\dot{V}o_2$peak despite higher work rates during verification tests could be rest time and intensity used. Nolan et al. (28) found that 105% maximal workload was more successful in $\dot{V}o_2$max confirmation than 115%, with 20-minute rest being time efficient and effective. Intensity chosen for verification testing affects the

---

Table 1

Peak measurements of performance and physiological variables during testing.*

Visit	Test	Time (s)	Heart rate (% APMHR)	RER	$\dot{V}o_2$ (ml·kg^{-1}·min^{-1})
1	Incremental	580 [562, 597]	97 [96, 99]	1.17 [1.14, 1.20]	50.0 [48.4, 51.5]
	Verification	160 [141, 179]	94 [92, 95]	1.11 [1.08, 1.14]	47.6 [46.0, 49.2]
2	Incremental	591 [574, 609]	97 [95, 98]	1.16 [1.13, 1.19]	49.9 [48.3, 51.5]
	Verification	164 [145, 182]	93 [92, 95]	1.09 [1.05, 1.12]	48.9 [47.3, 50.5]

*APMHR, age-predicted maximal heart rate (220 − age) (6); RER, respiratory exchange ratio; $\dot{V}o_2$, oxygen uptake.
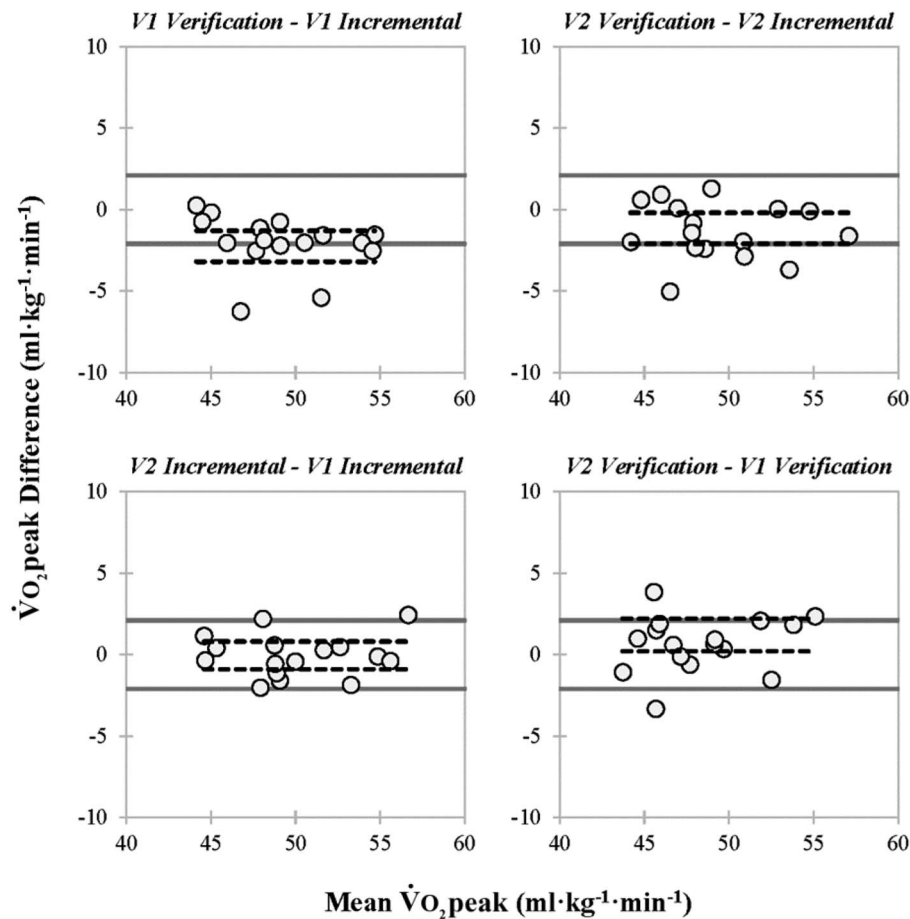
Figure 1. Bland-Altman plots of agreement between incremental and verification tests on visit 1 (V1) and visit 2 (V2). Black dashed lines, 90% confidence limits; dark gray lines, equivalence limits (± 2.1 ml·kg^{-1}·min^{-1}).

time to exhaustion, which may hinder the building of the slow component of V̇O₂ kinetics (30). Our volunteers lasted on average 163 seconds during verification testing. Poole and Jones (31) suggested 110% workload verification tests should last 180–540 seconds and test that do not reach this range are likely "extreme workloads." As long as volunteers are provided at least several minutes of recovery time, previous work suggests the exact rest period between incremental exercise test and the verification test may be less important (31). In a young and healthy population, rest periods between 5 and 60 minutes have been used and well tolerated (37). However, the lower-than-expected time to exhaustion during verification testing may be the product of both insufficient rest and excessive intensity (30) as well as the fasting state of the volunteers. Although an overnight fast does not alter pre-exercise muscle glycogen levels (13), pre-exercise glucose ingestion can increase time to exhaustion in high-intensity exercise (11). Balancing of sufficient rest and appropriate intensity for healthy individuals continues to elude researchers and should be evaluated on a per-cohort-basis (31,37).

Conversely, Scharhag-Rosenberger et al. (36) administered a treadmill incremental exercise test to 40 volunteers at 0.4% constant incline, then an initial 110% maximal velocity verification test 10 minutes later. Volunteers who were unable to exceed their incremental test V̇O₂peak during the initial verification test came back the next day for a follow-up 110% maximal velocity verification test. Those who exceeded their incremental test V̇O₂peak were instead given 10-minute rest before completing a

second verification test at 115% maximal velocity. Scharhag-Rosenberger et al. (36) saw 34 of 40 achieve "true" V̇O₂max within the initial verification with the same intensity and shorter rest than the current study. Moreover, 4 of the 6 who exceeded their incremental running V̇O₂peak with the initial verification test achieved "true" V̇O₂max verification with the 115% maximal velocity test, only 10 minutes after the initial verification test. Verification test at 110% maximal velocity on day 1 and day 2 was equivalent in V̇O₂peak despite day 2 verification lasting longer (126 $\pm$ 22 seconds; 162 $\pm$ 38 seconds) (36). Similarly, Rossiter et al. (33) used a 5-minute active recovery at 20 W after a ramp incremental test on a cycle ergometer followed by a verification step exercise at 5% higher work rate which lasted on average 88 seconds with a mean difference between ramp protocol and verification test of 31 ml·min^{-1} ($<1\%$).

The utility of a verification test to discern whether a "true" V̇O₂max is attained is dependent on a proper protocol that allows the individual to reach an equivalent or higher-than-previously-recorded V̇O₂peak, if possible (37). In the current study, volunteers reached V̇O₂peak measures that were not statistically equivalent despite a higher work rate. There were 6 instances in which a volunteer reached a higher V̇O₂ on their verification test versus the incremental test, but 5 occurred on the second visit (Figure 1). In addition, none of the verification tests exceeded the preceding incremental V̇O₂peak by ≥ 2.1 ml·kg^{-1}·min^{-1}; the limit we selected for statistical equivalence. The combination of rest period and intensity of verification testing used in the current

study does not seem to give active duty military populations the best opportunity to reach an equivalent V̇O₂peak to their preceding Astrand protocol.

Results from the current study suggest that repeating an incremental exercise test for familiarization or using a verification test to confirm V̇O₂max attainment may not be necessary in the active duty military population for general purposes. In some circumstances, however, the lowest margin of error is essential when determining V̇O₂max or when trying to find marginal changes in V̇O₂max. Although our study did not demonstrate a large benefit from familiarization, this approach may be suitable for researchers focused on obtaining the highest possible V̇O₂ measurements. From this perspective, additional testing provide volunteers with the opportunity to exceed their best performance, even if the difference is less than measurement precision or day-to-day variability. In individuals with exceptionally high aerobic fitness, additional V̇O₂max improvements from years of training would be minimal at best (21). For instances when only minute changes are expected, confirming a "true" V̇O₂max is paramount and extra familiarization or verification testing is warranted to create a low margin of error. In addition, identifying optimal secondary criteria and more precise cutoff values for V̇O₂max attainment in active duty military personnel could be valuable for researchers that cannot repeat incremental tests or use verification strategies.

Our study sample was sufficient for addressing the research aims of the present investigation but only included 1 female subject. Further recruitment of female soldier volunteers was unfortunately prevented by logistical, budgetary, and financial limitations imposed by the coronavirus disease 2019 (COVID-19) pandemic. Previous research has shown that sex does not influence discrepancies in V̇O₂max measurements between incremental and verification tests (3). In addition, elite male and female runners with similar performance levels had similar time to exhaustion at, and above, V̇O₂max (6). This suggests that sex alone is not determinant of the ability to reach a V̇O₂ plateau or perform a verification trial. Although our study had a comparable percentage of female soldiers to current U.S. Army combat position assessment (9), future studies need additional female volunteers to match the sex distribution of upcoming generations of American warfighters.

In conclusion, military researchers should expect the majority of soldiers can attain V̇O₂max by performing a modified Astrand treadmill running test. Additional verification and familiarization strategies may not be necessary in healthy active military personnel unless the lowest margin of error is essential. Verification testing may be more effective with a smaller work rate increase (<10%) or longer recovery time (>12 minutes) than used in this study.

Practical Applications

Active duty military personnel undergoing the modified Astrand protocol reliability attain their maximal oxygen uptake (V̇O₂max). Repeating incremental exercise test for familiarization purposes or V̇O₂max verification testing may not be necessary in active duty military cohorts unless the lowest margin of error is essential. A 10% increase in speed after a modified Astrand protocol for verification testing may be too intense in active duty military populations.

References

1. American College of Sports Medicine. Appendix D. In: *ACSM's Guidelines for Exercise Testing and Prescription*. Philadelphia, PA: Lippincott Williams & Wilkins, 2013. pp. 300–312.
2. Astorino TA. Alterations in V̇O₂max and the V̇O₂ plateau with manipulation of sampling interval. *Clin Physiol Funct Imaging* 29: 60–67, 2009.
3. Astorino TA, DeRevere J. Efficacy of constant load verification testing to confirm VO₂max attainment. *Clin Physiol Funct Imaging* 38: 703–709, 2018.
4. Astorino TA, Rietschel JC, Tam PA, et al. Reinvestigation of optimal duration of VO₂max testing. *J Exerc Physiol* 7: 1–8, 2004.
5. Bassett DR Jr, Howley ET. Limiting factors for maximum oxygen uptake and determinants of endurance performance. *Med Sci Sports Exerc* 32: 70–84, 2000.
6. Billat V, Beillot J, Jan J, Rochcongar P, Carre F. Gender effect on the relationship of time limit at 100% VO2max with other bioenergetic characteristics. *Med Sci Sports Exerc* 28: 1049–1055, 1996.
7. Buch R, Dysvik A, Kuvaas B, Säfvenbom R. The relationship between mastery orientation and maximal oxygen uptake among military cadets: The mediating role of intrinsic motivation. *Mil Behav Health* 4: 398–408, 2016.
8. Buchfuhrer MJ, Hansen JE, Robinson TE, et al. Optimizing the exercise protocol for cardiopulmonary assessment. *J Appl Physiol* 55: 1558–1564, 1983.
9. Department of the Army. *Annual Report on Progress of the Army in Integrating Women into Military Occupational Specialties and Units Recently Opened to Women*. In: Dot army. Washington, DC: Headquarters, Department of the Army, 2019. pp: 1–30.
10. Fox S, Haskell W. The exercise stress test: Needs for standardization. In: *Cardiology: Current Topics and Progress*. New York, NY: Academic Press, 1970. pp: 149–154.

11. Galloway SD, Lott MJ, Toulouse LC. Preexercise carbohydrate feeding and high-intensity exercise capacity: Effects of timing of intake and carbohydrate concentration. *Int J Sport Nutr Exerc Metab* 24: 258–266, 2014.

12. Green S, Askew C. V̇o₂peak is an acceptable estimate of cardiorespiratory fitness but not V̇o₂max. *J Appl Physiol* 125: 229–232, 2018.

13. Hearris MA, Hammond KM, Fell JM, Morton JP. Regulation of muscle glycogen metabolism during exercise: Implications for endurance performance and training adaptations. *Nutrients* 10: 1–21, 2018.

14. Hill A, Lupton H. Muscular exercise, lactic acid, and the supply and utilization of oxygen. *QJM* os-16: 135–171, 1923.

15. Kaminsky LA, Arena R, Myers J. Reference standards for cardiorespiratory fitness measured with cardiopulmonary exercise testing: Data from the Fitness Registry and the Importance of Exercise National Database. *Mayo Clin Proc* 90: 1515–1523, 2015.

16. Knapik J, Banderet L, Bahrke M, et al. Technical Report, specifically Technical Report No. T94-7. In: *Army Physical Fitness Test (APFT): Normative Data on 6022 Soldiers*. Natick, MA: U.S. Army Research Institute of Environmental Medicine, 1994.

17. Knapik JJ, Redmond JE, Grier TL, Sharp MA. Secular trends in the physical fitness of United States Army infantry units and infantry soldiers, 1976-2015. *Mil Med* 183: e414–e426, 2018.

18. Knapik JJ, Sharp MA, Darakjy S, et al. Temporal changes in the physical fitness of US Army recruits. *Sports Med* 36: 613–634, 2006.

19. Knapik JJ, Sharp MA, Steelman RA. Secular trends in the physical fitness of United States Army recruits on entry to service, 1975-2013. *J Strength Cond Res* 31: 2030–2052, 2017.

20. Lakens D. Equivalence tests: A practical primer for t tests, correlations, and meta-analyses. *Soc Psychol Personal Sci* 8: 355–362, 2017.

21. Legaz AA, Serrano EO, Jcasajús JM, Munguía DI. The changes in running performance and maximal oxygen uptake after long-term training in elite athletes. *J Sports Med Phys Fitness* 45: 435–440, 2005.

22. Mello RP, Murphy MM, Vogel JA. Relationship between a two mile run for time and maximal oxygen uptake. *J Strength Cond Res* 2: 9–12, 1988.

23. Meyer T, Scharhag J, Kindermann W. Peak oxygen uptake. *Z für Kardiologie* 94: 255–264, 2005.

24. 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* 34: 115–123, 2009.

25. Midgley AW, McNaughton LR, Polman R, Marchant D. Criteria for determination of maximal oxygen uptake. *Sports Med* 37: 1019–1028, 2007.

26. Myers J, Prakash M, Froelicher V, et al. Exercise capacity and mortality among men referred for exercise testing. *New Engl J Med* 346: 793–801, 2002.

27. Nakagawa S, Johnson PCD, Schielzeth H. The coefficient of determination R(2) and intra-class correlation coefficient from generalized linear mixed-effects models revisited and expanded. *J R Soc Interf* 14: 1–11, 2017.

28. Nolan P, Beaven M, Dalleck L. Comparison of intensities and rest periods for VO2max verification testing procedures. *Int J Sports Med* 35: 1024–1029, 2014.

29. Pollock ML, Bohannon RL, Cooper KH, et al. A comparative analysis of four protocols for maximal treadmill stress testing. *Am Heart J* 92: 39–46, 1976.

30. Poole DC, Jones AM. Oxygen uptake kinetics. *Compr Physiol* 2: 933–996, 2011.

31. Poole DC, Jones AM. Measurement of the maximum oxygen uptake Vo2max: Vo2peak is no longer acceptable. *J Appl Physiol* 122: 997–1002, 2017.

32. R Core Team. *R: A Language and Environment for Statistical Computing*. Vienna, Austria: R Foundation for Statistical Computing, 2014.

33. Rossiter H, Kowalchuk J, Whipp B. A test to establish maximum O2 uptake despite no plateau in the O2 uptake response to ramp incremental exercise. *J Applied Physiol* 100: 764–770, 2006.

34. Rusticus SA, Lovato CY. Applying tests of equivalence for multiple group comparisons: Demonstration of the confidence interval approach. *Pract Assess Res Eval* 16: 1–6, 2011.

35. Sandbakk SB, Nauman J, Zisko N, et al. Sedentary time, cardiorespiratory fitness, and cardiovascular risk factor clustering in older adults—the Generation 100 study. *Mayo Clinic Proc* 91:1525–1534, 2016.

36. 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* 36: 153–160, 2011.

37. Schaun GZ. The maximal oxygen uptake verification phase: A light at the end of the tunnel? *Sports Med Open* 3: 1–15, 2017.

38. Taylor HL, Buskirk E, Henschel A. Maximal oxygen intake as an objective measure of cardio-respiratory performance. *J Appl Physiol* 8: 73–80, 1955.

39. Weatherwax RM, Richardson TB, Beltz NM, Nolan PB, Dalleck L. Verification testing to confirm VO2max in altitude-residing, endurance-trained runners. *Int J Sports Med* 37: 525–530, 2016.

40. Witzki A, Rohde U, Nestler K, Leyk D. Achievement motivation in basic military task performance. *J Sci Med Sport* 20: S52–S52, 2017.