PREGLEDNI ZNANSTVENI RAD
REVIEWS

Hrvat. Športskomed. Vjesn. 2024; 39:5-13

CONTEMPORARY METHODS OF PREDICTING MAXIMUM OXYGEN UPTAKE: A BRIEF REVIEW

SUVREMENE METODE ZA PROCJENU MAKSIMALNOG PRIMITKA KISIKA: KRATKI PREGLED

Indy Man Kit Ho, Andro Matković[1], Branka R. Matković[2]
Technological and Higher Education Institute of Hong Kong, Department of Sports and Recreation, China,
[1]Clinical Department of Diagnostic and Interventional Radiology, Merkur University Hospital, Zagreb, Croatia,
[2]University of Zagreb Faculty of Kinesiology, Zagreb, Croatia
Correspondence: Andro Matković, andro.matkovic@gmail.com
DOI 10.69589/hsv.39.1.1

SAŽETAK

Kardiorespiratorna izdržljivost jedna je od najvažnijih fitness kvaliteta za sve populacije, uključujući zdrave pojedince, starije osobe, pacijente s kroničnim bolestima, rekreativne trkače i vrhunske sportaše. Potrošnja kisika u tjelesnim tkivima povećava se povećanjem aktivnosti ili intenziteta vježbanja, te količina kisika koja se potroši u minuti naziva se primitak kisika (VO2). Najveća količina kisika koju organizam može potrošiti u jednoj minuti naziva se maksimalni primitak kisika (VO2max).

U utvrđivanju VO2max-a, sportski znanstvenici obično provode izravno mjerenje pomoću inkrementnog protokola opterećenja na traci ili biciklu, a takav laboratorijski pristup mjerenju VO2max smatra se zlatnim standardom. Ipak, pristupačnost opreme i troškovi ispitivanja, kao i osoblje i stručnost koje treba uključiti, smatraju se sve faktorima koji test čine previše kompliciranim pa i skupim. Vezano uz sve navedeno procjena VO2maxa pomoću širokog raspona metoda s prihvatljivom pouzdanošću najprikladnija je alternativa.

Svrha ovog kratkog pregleda je kritička rasprava o uobičajenim metodama procjene, uključujući njihovu praktičnu primjenu, pouzdanost, valjanost i potencijalna ograničenja svake metode.

Ključne riječi: primitak kisika, procjena VO2max, testovi za procjenu VO2max

SUMMARY

Cardiorespiratory endurance is one of the most important fitness qualities for all populations including healthy individuals, the elderly, patients with chronic illness, recreational runners, and elite athletes. The uptake of oxygen by body tissues increases by increasing the activity or exercise intensity, also known as oxygen uptake (VO_2). When the VO_2 has reached the highest point that no additional oxygen can be further consumed by our cells, the maximum oxygen uptake (VO_2max) is achieved.

In assessing VO_2max, sports scientists commonly conduct the direct measure using the incremental graded testing protocol on a treadmill or bike, and such laboratory-based VO_2max is also regarded as the gold standard. Nevertheless, equipment accessibility and the testing cost as well as the personnel or expertise to be involved are all considered factors that make the test over-complicated and cumbersome. In this regard, the prediction of VO_2max using a wide range of methods with acceptable testing accuracy and time cost will be the most feasible alternative.

Therefore, the purpose of this brief review is to critically discuss the common types of prediction methods including their practical applications, the reliabilities, validities, and potential limitations of each method.

Keywords: Oxygen consumption, Prediction of VO2max, Exercise testing

Cardiorespiratory endurance is an essential fitness quality for all populations, including healthy individuals, the elderly, patients with chronic illness, recreational runners, and elite athletes (2,9,40). In daily life, the body produces and supplies the required energy via the oxidative pathway (i.e., the aerobic energy system).

During daily or sports activities, the oxygen is delivered via the cardiovascular and respiratory systems and finally consumed by our target working tissues, skeletal muscles. Oxygen uptake by body tissues increases by increasing the activity or exercise intensity, also known as oxygen uptake (VO2). When the VO2 has reached the highest point, and no additional oxygen can be further consumed by our cells, the maximum oxygen uptake (VO2max) is achieved (2,9,40). Even when athletes are required to exercise at an intensity higher than VO2max the adenosine triphosphate (ATP) will be synthesized via the anaerobic system, including the anaerobic glycolysis and the ATP-creatine phosphate system (ATP-CrP), the aerobic system is still highly demanded for supplying the ATP and facilitate the recovery of CrP via oxidative phosphorylation (50). Therefore, measuring the VO2max to understand competitive athletes' health status and performance potential is of great interest.

In assessing VO2max, sports scientists commonly conduct the direct measure using the incremental graded testing protocol on a treadmill or bike, and such laboratory-based VO2max is also regarded as the gold standard. Nevertheless, equipment accessibility, the testing cost, and the personnel or expertise to be involved or testing anyone not safe to push to exhaustion are all considered factors that make the test over-complicated and cumbersome. It is also infeasible and time-consuming for coaches to adopt direct VO2max assessment for teams or multiple athletes. The most feasible alternative is predicting VO2max using various methods with acceptable testing accuracy and time cost (35,40). Moreover, it was argued that the lab-based direct measures on VO2 max in a controlled environment (i.e., terrain, weather, wind, movement patterns) needed to reflect the actual athletic performance during the game or tournament conditions (16). Therefore, this brief review aims to critically discuss the common types of prediction methods, including their practical applications, the reliabilities, validities, and potential limitations of each method.

CONTINUOUS ACTIVITIES FOR VO2MAX PREDICTION

Most indirect VO2max estimation methods are conducted in a field setting using submaximal activities to lower exhaustion levels and risks. At the same time, some tests predict oxygen consumption based on the athletes' exercise heart rate, distance covered, or actual performance in a time trial. The type of exercise (non-impact vs. impact), such as walking, cycling, running, and swimming, could produce different testing results on the same individual.

Cycle based: Åstrand-Rhyming cycle ergometer test

The Åstrand-Rhyming cycle ergometer test is one of the oldest tests used to predict VO2max (5), but it is still widely used. It is based on the almost linear relationship between heart rate and metabolism, especially in the 120 to 170 heartbeats/min range. The authors reported a difference of approximately ±10% between predicted and directly measured VO2max.

The test is performed on a bicycle ergometer for 5 to 6 minutes. Subjects are subjected to continuous exercise, which is expected to increase heart rate in the range of 120 to 170 beats/min at steady state. The value taken for the prediction is the heart rate measured at the end of the fifth (eventually sixth) minute and the load of the cycle ergometer.

If the heart rate does not stabilize between the fourth and fifth minute, the test is extended for another minute. Also, if the heart rate rises above 170 bpm at any time during the test, the test has to be interrupted and repeated after rest at a lower load intensity. When the heart rate does not reach 120 bpm during the first or second minute, the load should be increased, and then the test is extended to the sixth minute.

It is essential to say that there are different nomograms for men and women and correction factors for children and those older than 35.

Since the introduction of this test, its validity has been tested by many authors (17,21,36,42). Upon testing the validity and reliability of the Astrand nomogram, Macsween (27) concluded that extrapolation of submaximal heart rate could be used confidently for clinical monitoring and research purposes in evaluating VO2 max. The other studies also confirmed that the Åstrand test should be considered highly valid and feasible.

Running based: Cooper's 12-minute and 1.5-mile run

Running-based field tests are the most popular to assess the cardiorespiratory endurance of a wide range of active individuals or trained athletes. The Cooper's 12-minute and 1.5-mile runs are the two standard methods for such purposes. The Cooper's 12-minute run can be performed on an outdoor track or a motorized treadmill using 1% inclination. By simply measuring the completion time and using the formula:

$$\text{VO2max} = (22.351 \times \text{kilometers}) - 11.288$$

the VO2max can be easily estimated. Cooper (15) reported a test-retest coefficient of reliability correlation of 0.98, while the correlation coefficient between the 12-minute run and the direct measure from the treadmill was 0.90. The more recent study conducted by Penry et al. (37) also showed comparable criterion-based validity ($r = 0.87$) using sixty healthy young adults (33 women and 28 men between 18 and 33 years) and good reliability using G-study. Similarly, the 1.5-mile run can be performed on an outdoor track or indoor treadmill, whereas the VO2max can be estimated using the formula.

VO2max = 91.736 – (0.1656 x kg) – (2.767 x completion time in min) for males, and

VO2max = 88.020 – (0.1656 x kg) – (2.767 x completion time in min) for females (32).

The validity using corrected mean r at the population level (rp = 0.79, 0.73–0.85) is acceptable compared to the direct measure gold standard. Moreover, both 1.5-mile and 12-minute runs were shown to have significantly higher criterion-related validity than other tests such as 3 miles, 1-mile, 1/2-mile, 600 yards, 1/4 mile, 15- or 6-minute run tests ($p<0.05$) (32). Therefore, these two simple running-based methods in estimating VO2max can be accepted as reliable and valid alternatives when the laboratory is not accessible, and the graded exercise protocol is not a feasible option. However, Cooper's 12-minute run requires athletes to pace themselves to adopt a maximum steady and sustainable speed throughout the test. Therefore, inexperienced runners without previous testing experience may have difficulty accurately selecting the proper running tempo to optimize the testing performance.

Running-based: Time-trial-based predictions using the VDOT formula

Although VO2max is recognized as the best indicator to reflect individual cardiorespiratory endurance, the magnitude of the improvement after a certain training period highly depends on the fitness level of the runners. For example, runners with lower baseline fitness levels improve more substantially than those already well-trained. Moreover, it was found that VO2 max may not be the most critical parameter to determine the winners in long-distance running events (e.g., 10 km). Therefore, it is believed that both VO2max and the lactate threshold (LT) have contributed to the overall endurance performance. Nevertheless, the concurrent direct measures on VO2max and LT in the well-equipped laboratory are expensive and complicated.

Moreover, the performance (VO2max) observed in the controlled environment (i.e., flat motorized treadmill with constant temperature, humidity, and wind) cannot truly reflect the actual performance during the competition. To help coaches and runners overcome such assessment problems, a well-known running coach and scientist, Jack Daniels, has developed the VDOT running calculator (16). The VDOT is similar to the concept of VO2max, whereas the VDOT value is obtained via practical time trials to estimate. Despite the high popularity and recognition of the VDOT in the running industry, the validity of VDOT to predict VO2max was not widely studied until Scudamore et al. (40) made an initial attempt at it recently. When the indoor 5 km time trials were used to perform VDOT calculation and compare to the laboratory VO2max, it was found that VDOT underestimated VO2max in both NCAA Division 1 track athletes ($p<0.01$; effect size: $d = 1.75$) and recreational runners ($p < 0.01$; effect size: $d = 3.44$). Although the authors recommended that athletes and coaches interpret VDOT findings cautiously due to underestimation, the VO2max and VDOT could be used for different purposes. When benchmarking the cardiorespiratory fitness of athletes to others is the priority, the VO2max should be a better choice since this value solely focuses on a particular fitness quality. In contrast, if the actual running performance is to be assessed for designing the tempo run or interval training pace, the VDOT can be a more feasible option to give meaningful values.

Walking based: Rockport 1-mile walking test

Apart from running, the walking test is another method widely used to assess the cardiovascular function of various populations. Due to the low-impact characteristic, the walking-based VO2max estimation is a safer option for deconditioned or inactive individuals. The Rockport one-mile walking test, developed by Kline et al.(24), requires participants to walk as fast as they can in the standard 400-meter running track for four laps (1609 meters or approximately 1.6 km). The regression formula estimating the VO2max was developed using the completion time, the heart rate determined after the walk, the age, gender as well as the body weight in pounds as follows (34):

VO2 max = 88.7688 + 8.8924 x (sex: F=0; M=1) – 0.0957 x (body weight in pounds) – 1.4537 x (walk time in minutes) – 0.1194 x (HR at completion)

The Rockport one-mile walking test showed excellent test-retest reliability in estimating VO2 max using intraclass correlation coefficient (ICC) (0.96 to 0.97) when assessing middle-aged and older women ages 50-69 years (19). A more recent study showed no significant difference between the predicted VO2max in the Rockport walking test and the direct VO2max assessment using graded treadmill exercise protocol ($p = 0.18$, $r = 0.82$) on active-duty Air Force males (aged 18 to 44). (47). Similarly, McSwegin et al. (34) also demonstrated exceptionally high reliability (ICCs: 0.91 to 0.97) using different VO2max estimation equations and good validity ($r = 0.8$ to 0.84), standard errors of the estimate (SEE) = 4.50 to 4.99) on high school individuals. However, it is noteworthy that the degree of similarity in terms of metabolic demand and biomechanical characteristics of the activities between the testing activity and actual performance is critical to determine the sensitivity of the assessment method in reflecting the true fitness capabilities of athletes. Therefore, VO2max and cardiorespiratory performance predicted from walking tests are not optimum choices for competitive runners or swimmers to reflect the relevant performance. In this regard, despite the excellent validity observed using the Rockport 1-mile walking test on active high school individuals, the results may only be used as a reference to depict the general health rather than the actual performance in high-intensity sports.

Stepping-based: Chester step test

Besides the running and walking tests, the step tests are also widely used to predict cardiovascular performance in the indoor setting with minimum equipment required (stopwatch, metronome, and bench/step board). The Chester step test is one of the few using the incremental protocol that individuals must complete a multistage (the tempo increases from 60 to 140 beats per minute in a total of 5 stages) stepping activities on a 30 cm bench (the height is adjusted according to the age and fitness background). It was shown to have acceptable reliability, but the criterion-based validity was questionable (11 to 19% error) (8). In contrast, another study conducted by the original author promoting the use of the Chester step test demonstrated a high correlation coefficient (r = 0.92, overall standard error of estimate (SEE): ±3.9 ml O2/kg/min, 5 to 15%) when compared to the incremental VO2max treadmill test. The test-retest reliability of this step test was reported to be good using the Bland and Altman method (the mean difference between repeated predicted measures was −0.7 ml O2/kg/min while the inter-day reliability was acceptable with the difference within 4.5 ml/kg/min (44). Due to the inconsistent validity reported, further studies on using the Chester step test to predict VO2max for different populations are warranted.

Stepping-based: The Queen's College step test

Apart from the Chester Step Test, the Queen's College and YMCA 3-minute step tests are also widely used to assess general populations' cardiovascular fitness. The Queen's College step test is performed on a stool of 41.3 cm in height for 3 minutes of stepping exercise at 24 cycles per minute. After completion, the individuals can measure their carotid pulse rate from 5 to 20 seconds of the recovery period (a total of 15 seconds), and the predicted VO2max can be gauged by:

VO2max (ml/kg/min) = 111.332 - (0.426 x heart rate in beats/min) (11,41).

The Queen's College step test showed no significant difference between the direct VO2max measure and the treadmill protocol (p = 0.10; mean difference: 0.46 ml/min/kg) in a sedentary male university. Similarly, the YMCA 3-minute step test requires individuals to perform 72 steps (24 steps per minute on a 30 cm bench and 3 minutes). One minute of post-exercise heart rate is recorded for calculating predicted VO2max.

VO2max = 70.597 – 0.246 x (age) + 0.077 x (height) – 0.222 x (weight) – 0.147 x (heart rate) for males, or

VO2max = 70.597 – 0.185 x (age) + 0.097 x (height) – 0.246 x (weight) – 0.122 x (heart rate) for females.

It was shown to have good criterion-based validity using a correlation coefficient (r = 0.80, p < 0.01) (22). To conclude, using one of the step tests mentioned above to predict VO2max for active and inactive individuals is a feasible choice, especially when the outdoor track or treadmill is inaccessible. Moreover, these can be good alternatives when individuals have difficulties or limitations in walking or running tasks. Nevertheless, practicing trials is recommended, especially for those who are not familiar with the tempo given by the metronome.

Swimming-based: 12-minute Swim Test

Like the dryland performance assessment, the laboratory VO2max direct measure is an expensive option many club swimmers may need help accessing through the well-equipped testing facilities. Moreover, lab-based VO2max measures do not allow testing for a large group of athletes simultaneously; therefore, it is very time-consuming. The 12-minute swim test is similar to the Cooper's 12-minute run. (12). However, compared to the treadmill running or tethered swimming as the criterion-related validity, only moderately low correlation coefficients were observed (r = 0.38 to 0.4; standard errors of estimate = 5.1 to 5.7 ml/kg/min). Therefore, the authors concluded that it is not suitable to be the alternative to the traditional 12-minute run, but it is still an accurate and sensitive tool to help in classifying the aerobic performance of swimmers. These data were in close agreement with the results of the study performed a year later on female swimmers (13). The authors concluded that the 12-minute swim had relatively low validity and is not an equally valid alternative to the 12-minute run in young adult female recreational swimmers. Nonetheless, only moderate test-retest reliability (r = 0.66) and moderately low validity (r = 0.47) were observed in another more recent study when tested on high school swimmers (aged 13 to 17) (20). In this regard, the usefulness of adopting such a 12-minute swim test for predicting VO2max and performance is highly questionable.

Dancing-based: Dance-Specific Aerobic Fitness Test (DAFT)

Unlike other competitive sports such as soccer or marathon, athletes always push themselves to the limit during the game; on the other hand, dancers have the intensity determined by the preset choreography and their extremely heavy skill-related component in this physical activity. Due to the poor movement and metabolic specificity between the dance moves and testing activities (e.g., treadmill running), the value of VO2max yielded from direct measure via maximal treadmill exercise is highly questionable (45,48). In this regard, Wyon et al. (49) have developed a novel test for dancers called the Dance

Specific Aerobic Fitness Test (DAFT). The 5-stage test requires dancers to perform the preset dance-specific drills for 4 minutes per stage, and the tempo increases from 68 bpm at stage one to 108 bpm at the final stage. Significant differences in VO2 and heart rate between stages were shown, and the correlation between heart rate and VO2 was high (r = 0.94; SEE = 4.51). Good reliability in terms of coefficient of variation (CV) was also observed between trials (1.4 to 6.0). By using the simple linear regression, the formula for predicting VO2 is obtained:

VO2 = 0.4234 x (heart rate in beats per minute) – 27.829,

and the HR-VO2 relationship in this regression is strong as well (r = 0.91; SEE = 5.6 ml/kg/min). Similarly, the multistage ballet-specific aerobic fitness test was developed with comparable test-retest reliability and validity (45). However, it is worth noting that these studies did not push the dancers to the limit to estimate the VO2max or VO2peak. To estimate the VO2max of a dancer using these multistage dance-specific tests, the estimation of individual maximum heart rate (HRmax) is needed, and the actual difference between the true VO2max and the predicted VO2max using such regression formulas was not reported.

INTERMITTENT ACTIVITIES FOR VO2MAX PREDICTION

Although many running-based or stepping-based field tests can predict VO2max with good reliability and validity, due to the extremely high exercise intensity, intermittent running or sprinting in nature with frequent starts and stops, and the high involvement of anaerobic energy system, these methods may not truly reflect the potential and competence of trained athletes in sports with specific needs such as soccer, rugby, tennis, badminton, and basketball. Despite the high demand for explosive moves (sprinting, change of direction, jumping) and the demand on fast anaerobic energy systems (i.e., anaerobic glycolysis and creatine phosphate), the efficiency of oxidative aerobic system is still vital for both recovery and the replenishment of the phosphocreatine via the oxidative phosphorylation process (50). Therefore, instead of continuous walking or running-based testing methods, testing protocols with repeated intermittent runs or sprints provide more realistic field-based VO2max prediction for these athletes.

YoYo-Intermittent Recovery 1 Test

In this regard, Yo-Yo test variants were suggested and widely used to estimate the VO2max (39). The test was developed based on the maximal multistage 20-m shuttle-run test modified by an active recovery period (6,26). Conflicting findings were reported for the criterion-based validity of YoYo-Intermittent Recovery 1 (YYIR1) in predicting the maximum oxygen consumption compared to the laboratory VO2max. Acceptable validity was found when the YYIR1 test was applied to elite soccer players or recreationally active individuals (r = 0.71 to 0.87). At the same time, another study did not show a significant correlation between the directly measured VO2max from the laboratory and the predicted values from YYIR1 (39).

Similarly, a more recent study performed the YYIR1 test and retest on two separate sessions and another independent maximal performance running test with time-series analyses of gas exchange parameters. Good test-retest repeatability (Coefficient Variation = 6.63; ICC = 0.86) was observed. However, a significantly higher VO2peak during the first YYIR1 test than the retest (8.81 ± 5.6%) was reported; meanwhile, only a weak correlation coefficient (R2 = 0.28, p = .115) in terms of the VO2peak between YYIR1 and laboratory test was found (39). The authors concluded that YYIR1 was not suitable for estimating the VO2 peak. One of the possible explanations for such significant discrepancies in different studies can be the individual differences in the actual contribution of the aerobic and anaerobic energy systems to YYIR1 test performance. When performing the YYIR1 test, the majority of the time is supported by the aerobic system for elite athletes. At the same time, those low to moderately fit individuals heavily rely on the anaerobic contribution in most testing stages. Such a proposed explanation was echoed by another study in which low accuracy of VO2max prediction was observed in female soccer players (30).

YoYo-Intermittent Endurance 2 Test

The YoYo-Intermittent Endurance 2 (YYIE2) test starts at a higher running sprint (13 km/hr instead of 10 km/hr) when compared to the YYIR1, and the YYIE2 is more widely used for trained athletes. Similar to the YYIR1 test, conflicting findings regarding the inconsistent validity results (r = 0.43 to 0.75) were reported (39). The CVs ranged from 7.1% to 9.6% when the test was administered on trained soccer players but were higher (CV = 12.7%) when the tested participants were recreationally active men only. Similar to the YYIR1 test, when these YoYo test variants were adopted in schoolboys (aged 9 to 16), the CVs decreased with increasing age from 11.1 to 8.5%. The YYIE2 is more reliable when used in trained or elite individuals at least 16 years old. For the application, given that the true VO2max value is not the only determinant of the real game performance for many sports (except those endurance-based such as a marathon) and the laboratory-based VO2max is performed in a controlled environment using continuous treadmill run without the demand on acceleration, deceleration and the change of direction, scientists and coaches should choose YYIE2 to obtain the field aerobic performance to have a more realistic picture about the physical fitness quality of athletes when the schedule is tight and the testing time is limited. No matter which YoYo test is used, the equation for predicting VO2max is as follows (6):

- YYIR1 test: VO2max (mL/kg/min) = IR1 distance (m) × 0.0084 + 36.4

- YYIE2 test: VO2max (mL/kg/min) = IR2 distance (m) × 0.0136 + 45.3

Bangsbo and his coworkers (6) analyzed many different studies investigating Yo-Yo tests. They concluded that these tests can adequately discriminate players' performance at different competitive levels, between playing positions, and after periods of various training types. They suggested that sports characterized by intermittent exercise can use Yo-Yo tests to determine an athlete's ability to perform intense intermittent exercise.

20-m Shuttle Run Test

Apart from the YYIR1 and YYIE2 tests, the 20-m shuttle run, known as "PACER" or "Multistage fitness test," is one of the most commonly used field tests for estimating cardiorespiratory endurance. Similar to the Yo-Yo test variants, it requires minimal equipment and can be used to test many individuals simultaneously. It consists of one-minute stages of continuous, incremental speed running with speed starting from 8.5 km/h and increasing by 0.5 km/h per minute (32). Compared to the laboratory VO2max, the criterion-related validity was shown to have a moderate to high mean correlation coefficient (rp with the overall weighted mean of r corrected for sampling error and measurement error: 0.66 to 0.95). Both males and females were shown to have similar criterion-related validity for estimating VO2max. At the same time, the mean correlation coefficient was only moderate when used in children and moderate-to-high for adults. Regardless of the level of VO2max of the participants, the 20-m shuttle run test had a moderate to high criterion-related validity. These findings observed by Mayorga-Vega et al. (32) were comparable to another study conducted by Matsuzaka et al. (31). They proposed that the chronological age of children is a significant predictor of VO2 max but not in adults. This would lead to a lower validity value in children. On the other hand, children were probably less willing to endure the uncomfortable perceived exertion with strenuous effort when exhausted. If they were less motivated, the validity would be affected as well.

NON-EXERCISE-BASED PREDICTIVE EQUATIONS

Prediction of VO2max without exercise is a great advantage when it is needed for assessment in persons whose health condition does not allow more extraordinary efforts or if there are problems with the testing resources and equipment. Many different equations are described and validated in the literature, and they are easy to administer. These equations are often based on basic quantitative measures like body height, body mass percentage, body fat, resting heart rate, age, gender, or physical activity status.

Wang and his coauthors (47) provided a very comprehensive study in 2019. Reviewing the literature, they tried to summarize existing non-exercise-based VO2max prediction models and determine their application value. PubMed search yielded 60 equations, some of which were developed for both sexes (20), some of them only for women (21), and some of them only for men (19). After cross-validation of proposed equations, they concluded that non-exercise-based prediction models are practical and viable, especially for less fit individuals.

Besides field-based exercise testing, sports scientists have successfully derived a novel VO2max prediction equation by modifying previously published VO2max equations using the constant error values and the predicted residual sum of squares statistic and cross-validating the modified equations to identify the accuracy for estimating VO2max in aerobically trained men (28). The non-exercise-based VO2max prediction equation is:

VO2max (ml/min) = 27.387 x (weight in kg) + 26.634 x (height in cm) – 27.572 x (age in years) + 26.161 x (hours per week of training) + 114.904 x (intensity of training using the Borg 6-20 scale) + 506.752 x (natural log of years of training) – 4609.791

with r = 0.82, adjusted R2 value = 0.65, and SEE = 378 ml/min.

The % of total error remained 10%, and therefore, the author concluded that this non-exercise-based equation should be recommended for estimating VO2max (28).

The same authors prepared the equation for aerobically trained women (29):

$\dot{V}$O2max (mL·min−1) = 18.528 x (weight in kg) + 11.993 x (height in cm) − 17.197 x (age in yr) + 23.522 x (h·wk−1 of training) + 62.118 x (intensity of training using the Borg 6–20) + 278.262 x (natural log of years of training) − 1375.878

(R = 0.83, R2 adjusted = 0.67, and SEE = 259 mL·min−1).

Although the equations are valid for aerobically trained men and women, they have yet to be widely used by coaches or in recent literature over the past few years. One of the possible issues is the questionable feasibility of generalizing these equations to different populations and athletes with diversified fitness backgrounds. As many field tests are available and have already been proven valid but not time-consuming, coaches are still recommended to choose a more sport-specific field-based exercise testing to predict the VO2max for their athletes instead of "guessing" their fitness qualities from the formula without observing their actual performance.

PREDICTIONS WITH WEARABLE DEVICES

Recently, the rapid growth of wearable technology in training and fitness monitoring has drawn tremendous attention from sports scientists and coaches. Wrist devices such as heart rate watches are economical options for most runners, and most of these products also provide functions such as VO2max prediction. Given the large populations and most runners relying on the data from these wearable devices to modulate their training program, a considerable amount of research in this regard has been conducted to investigate the accuracy and validity of the measures(3,14,18,23,25,43,46). A recent systematic review and meta-analysis (35) has reviewed 14 validation studies and revealed that wearable devices using resting condition information in their algorithms to predict VO2max would lead to overestimation (bias = 2.17 ml/kg/min; limits of agreement = -13.07 to 17.41 ml/kg/min) whereas the exercise-based algorithms demonstrated a lower systematic and random error (bias = -0.09 ml/kg/min; limits of agreement = -9.92 to 9.74 ml/kg/min). Despite the limited findings on reliability (only three included studies of the systematic review), good test-retest reliability of VO2max prediction using wearable devices was reported (ICC > 0.90).

Different methodologies are adopted in estimating VO2max for the brand of wearable devices. Polar has adopted HR, HR variability, gender, age, body weight, height, and self-reported physical activity to predict VO2max. The systematic error was low (2.17 ml/kg/min), but interestingly, the random error was extensive (30.48 ml/kg/min). In contrast, Fitbit and Garmin adopt the First-beat Technologies in the VO2max prediction. The calculation requires entering personal information such as age, collecting the HR and exercise speed by performing an exercise test, segmenting the HR into several zones, evaluating the reliability of segments, and applying linear and nonlinear relationships between HR and speed to estimate VO2max. The systematic error was very low (0.09 ml/kg/min), while the random error was still considerable (9.83 ml/kg/min).

To conclude, runners and coaches should rely on the exercise-based algorithm to estimate the VO2max as this can greatly reduce systematic bias and random error. For sedentary or deconditioned individuals for whom exercise testing is not feasible, the estimated VO2max yielded from rest-based algorithms should be interpreted with caution.

PREDICTION USING MACHINE LEARNING ALGORITHMS

With the fast growth and advancement of technology, artificial intelligence techniques such as machine learning (ML) algorithms are commonly used to perform predictive analytics, including classification and regression tasks. Similar to the multiple linear regression model, By combining the predictive variables such as speed and heart rate from previous submaximal testing and the selected ML algorithms, the estimation of VO2max is possible, especially for those who are not suitable to participate in physical activity (1). A recent review has shown that several ML models were used for predicting VO2max, including the support vector machine (SVM) with Relief-F, multi-layer perceptron (MLP), decision tree (DT), artificial neural network (ANN), generalized regression neural network (GRNN) and random forest (4). Models were typically built using data obtained from exercise, non-exercise, or hybrid methods. The model performance was determined mainly by the R values (ranging from 0.38 to 0.97), SEE (ranging from 3.34 to 10.67 ml/kg/min), and root mean square error (RMSE) (ranging from 2.91 to 4.78). Among these models, the author concluded that ANN is the most accurate (R = 0.91 to 0.97; SEE = 3.34 ml/kg/min) for VO2max prediction.

Interestingly, ANN was developed with inspiration from the human neuron structure. The ANN simulates the neurons with several interconnected layers, forming a network. During the machine learning process, the nodes (or neurons) are associated with certain weights such that these weights are adjusted during the training according to the training dataset. With the inter-connected layers, ANNs provide good tolerance for fault when one or more specific components of the neural network are lost or missing. Beltrame et al. (7) applied ANN using the predictors of treadmill speed, slope, gender, exercise time, heart rate, and body mass index. ML performance and reliability can be further enhanced with more data (the sample number and heterogeneous group of participants) feeding into the model for training purposes. However, since most of the studies included in the review mainly focused on only healthy individuals or college-aged students, the accuracy of these models in predicting VO2max in other populations, such as elite athletes, the elderly, and disabled people, is still unclear.

CONCLUSION

As there are pros and cons in each of the VO2max estimation methods, the athletes and coaches should take all the factors above, including the reliability, validity, risk of the test, the value in reflecting the actual performance, as well as the time and money required into consideration before choosing the VO2max prediction tool. As submaximal exercise, wearable technology, and machine learning algorithms are more popular and frequently adopted by sports coaches, future studies are highly recommended to develop an extensive database of athletic profiles using these testing methods.

References

1. Abut F, Akay MF. Machine learning and statistical methods for the prediction of maximal oxygen uptake: recent advances. Med Devices (Auckl). 2015;8:369-79.
2. Alzamer H, Abuhmed T, Hamad K. A short review on the machine learning-guided oxygen uptake prediction for sport science applications. Electronics. 2021 Aug 13;10(16):1956.
3. Anderson JC, Chisenall T, Tolbert B, Ruffner J, Whitehead PN, Conners RT. Validating the commercially available Garmin Fenix 5x wrist-worn optical sensor for aerobic capacity. Int J Innov Educ Res. 2019;7:147–58.
4. Ashfaq A, Cronin N, Müller P. Recent advances in machine learning for maximal oxygen uptake (VO2 max) prediction: A review. Inform Med Unlocked. 2022;28:100863.
5. Astrand PO, Ryhming I. A nomogram for calculation of aerobic capacity (physical fitness) from pulse rate during sub-maximal work. J Appl Physiol. 1954;7(2):218-21.
6. Bangsbo J, Iaia FM, Krustrup P. The Yo-Yo intermittent recovery test: a useful tool for evaluation of physical performance in intermittent sports. Sports Med. 2008;38:37-51.
7. Beltrame T, Amelard R, Villar R, Shafiee MJ, Wong A, Hughson RL. Estimating oxygen uptake and energy expenditure during treadmill walking by neural network analysis of easy-to-obtain inputs. J Appl Physiol (1985). 2016;121(5):1226-33.
8. Buckley JP, Sim J, Eston RG, Hession R, Fox R. Reliability and validity of measures taken during the Chester step test to predict aerobic power and to prescribe aerobic exercise. Br J Sports Med. 2004 Apr;38(2):197-205.
9. Buttar KK, Saboo N, Kacker S. A review: Maximal oxygen uptake (VO2max) and its estimation methods. Int J Phys Educ Sports Health. 2019;6(6):24-32.
10. Carrier B, Creer A, Williams LR, Holmes TM, Jolley BD, Dahl S, et al. Validation of Garmin Fenix 3 HR fitness tracker biomechanics and metabolics (VO2max). J Meas Phys Behav. 2020;3:331–7.
11. Chatterjee S, Chatterjee P, Mukherjee PS, Bandyopadhyay A. Validity of Queen's College step test for use with young Indian men. Br J Sports Med. 2004;38(3):289-91.
12. Conley DS, Cureton KJ, Dengel DR, Weyand PG. Validation of the 12-min swim as a field test of peak aerobic power in young men. Med Sci Sports Exerc. 1991;23(6):766-73.
13. Conley DS, Cureton KJ, Hinson BT, Higbie EJ, Weyand PG. Validation of the 12-minute swim as a field test of peak aerobic power in young women. Res Q Exerc Sport. 1992.;63(2):153-61.
14. Cooper KD, Shafer AB. Validity and reliability of the Polar A300's fitness test feature to predict VO2max. Int J Exerc Sci. 2019;12:393–401.
15. Cooper KH. A means of assessing maximal oxygen intake. Correlation between field and treadmill testing. JAMA. 1968;203(3):201-4.
16. Daniels J. Daniels' Running Formula. Champaign, IL: Human Kinetics, Inc. 2014.
17. Ekblom B, Engstrom LM, Ekblom O. Secular trends of physical fitness in Swedish adults. Scand J Med Sci Sports, 2007;17:267–73.
18. Freeberg KA, Baughman BR, Vickey T, Sullivan JA, Sawyer BJ. Assessing the ability of the Fitbit Charge 2 to accurately predict VO2max. Mhealth. 2019;5:39.
19. Hageman PA, Walker SN, Pullen CH, Pellerito P. Test-retest reliability of the Rockport Fitness Walking Test and other fitness measures in women ages 50–69 years. J Geriatr Phys Ther. 2001; 24(2): 7–11.
20. Huse D, Patterson P, Nichols J. The validity and reliability of the 12-minute swim test in male swimmers ages 13-17. MPEES. 2000;4(1):45-55.
21. Jessup GT, Riggs CE, Lambert J, Miller WD. The effect of pedalling speed on the validity of the Astrand-Rhyming aerobic work capacity test. J Sports Med Phys Fitness. 1977; 17(4):367-71.
22. Kieu NTV, Jung SJ, Shin SW, Jung HW, Jung ES, Won YH, et al. The validity of the YMCA 3-minute step test for estimating maximal oxygen uptake in healthy Korean and Vietnamese adults. J Lifestyle Med. 2020;10(1):21-29.
23. Klepin K, Wing D, Higgins M, Nichols J, Godino JG. Validity of cardiorespiratory fitness measured with Fitbit compared to VO2max. Med Sci Sports Exerc. 2019;51:2251–6.
24. Kline GM, Porcari JP, Hintermeister R, et al. Estimation of VO2max from one-mile track walk, gender, age, and body weight. Med Sci Sports Exerc. 1987;19(3):253-59.
25. Kraft GL, Roberts RA. Validation of the Garmin Forerunner 920XT Fitness Watch VO2peak test. Int J Innov Educ Res. 2017;5:63–9.
26. Leger LA, Mercier D, Gadoury C, Lambert J. The multistage 20-metre shuttle run test for aerobic fitness. J Sports Sci. 1988;6(2):93-101.
27. Macsween A. The reliability and validity of the Astrand nomogram and linear extrapolation for deriving VO2max from submaximal exercise data. J Sports Med Phys Fitness. 2001;41(3):312-7.
28. Malek MH, Housh TJ, Berger DE, Coburn JW, Beck TW. A new non-exercise-based Vo2max prediction equation for aerobically trained men. J Strength Cond Res. 2005;19(3):559-65.
29. Malek MH, Housh TJ, Berger DE, Coburn JW, Beck TW. A new nonexercise-based VO2max equation for aerobically trained females. Med Sci Sports Exerc. 2004;195:1804–10.

30. Martínez-Lagunas V, Hartmann U. Validity of the Yo-Yo Intermittent Recovery Test Level 1 for direct measurement or indirect estimation of maximal oxygen uptake in female soccer players. Int J Sports Physiol Perform. 2014;9(5):825-31.

31. Matsuzaka A, Takahashi Y, Yamazoe M, Kumakura N, Ikeda A, Wilk B, Bar-Or O. Validity of the multistage 20-m shuttle-run test for Japanese children, adolescents, and adults. Pediatr Exerc Sci. 2004;16(2):113-25.

32. Mayorga-Vega D, Aguilar-Soto P, Viciana J. Criterion-related validity of the 20-m shuttle run test for estimating cardiorespiratory fitness: A meta-analysis. J Sports Sci Med. 2015;14(3):536-47.

33. Mayorga-Vega D, Bocanegra-Parrilla R, Ornelas M, Viciana J. Criterion-related validity of the distance- and time-based Walk/Run Field Tests for estimating cardiorespiratory fitness: A systematic review and meta-analysis. PLoS One. 2016;11(3):e0151671.

34. McSwegin PJ, Plowman SA, Wolff GM, Guttenberg GL. The validity of a one-mile walk test for high school-age individuals. Meas Phys Educ Exerc Sci. 1998;2(1):47-63.

35. Molina-Garcia P, Notbohm HL, Schumann M, Argent R, Hetherington-Rauth M, Stang J, et al. Validity of estimating the maximal oxygen consumption by consumer wearables: A systematic review with meta-analysis and expert statement of the INTERLIVE Network. Sports Med. 2022;52(7):1577-97.

36. Nordgren B, Fridén C, Jansson E, Österlund T, Grooten WJ, Opava CH, Rickenlund A. Criterion validation of two submaximal aerobic fitness tests, the self-monitoring Fox-walk test and the Åstrand cycle test in people with rheumatoid arthritis. BMC Musculoskelet Disord. 2014;15:305.

37. Penry JT, Wilcox AR, Yun J. Validity and reliability analysis of Cooper's 12-minute run and the multistage shuttle run in healthy adults. J Strength Cond Res. 2011;25(3):597-605.

38. Schmitz B, Pfeifer C, Kreitz K, Borowski M, Faldum A, Brand SM. The Yo-Yo Intermittent tests: A systematic review and structured compendium of test results. Front Physiol. 2018;9:870.

39. Schmitz B, Pfeifer C, Thorwesten L, Krüger M, Klose A, Brand SM. Yo-Yo Intermittent Recovery Level 1 Test for estimation of peak oxygen uptake: Use without restriction? Res Q Exerc Sport. 2020;91(3):478-87.

40. Scudamore EM, Barry VW, Coons JM. An evaluation of Time-Trial-Based Predictions of VO2max and recommended training paces for collegiate and recreational runners. J Strength Cond Res. 2018;32(4):1137-43.

41. Shamsi MM, Alinejad HA, Ghaderi M, Badrabadi KT. Queen's College step test predicted VO2max: The effect of stature. Ann Biological Res. 2011;2(6):371-7.

42. Siconolfi SF, Cullinane EM, Carleton RA, Thompson PD. Assessing VO2max in epidemiologic studies: modification of the Astrand-Rhyming test. Med Sci Sports Exerc. 1982; 14(5):335–8.

43. Snyder NC, Willoughby CA, Smith BK. Comparison of the Polar V800 and the Garmin Forerunner 230 to predict $\dot{V}O_2$max. J Strength Cond Res. 2021; 35(5):1403-9.

44. Sykes K, Roberts A. The Chester step test—a simple yet effective tool for the prediction of aerobic capacity. Physiotherapy. 2004;90(4):183-8.

45. Twitchett E, Nevill A, Angioi M, Koutedakis Y, Wyon M. Development, validity, and reliability of a ballet-specific aerobic fitness test. J Dance Med Sci. 2011;15(3):123-7.

46. Wagner M, Engel F, Klier K, Klughardt S, Wallner F, Wieczorek A. On the reliability of wearable devices using the example of a premium multisport smartwatch. Ger J Exerc Sport Res. 2021;51:49-62.

47. Wang Y, Chen S, Lavie CJ, Zhang J, Sui X. An overview of non-exercise estimated cardiorespiratory fitness: estimation equations, cross-validation and application. J Sci Sport Exercise. 2019;1:38-53.

48. Weiglein L, Herrick J, Kirk S, Kirk EP. The 1-mile walk test is a valid predictor of VO(2max) and is a reliable alternative fitness test to the 1.5-mile run in U.S. Air Force males. Mil Med. 2011;176(6):669-73.

49. Wyon M, Redding E, Abt G, Head A, Sharp NC. Development, reliability, and validity of a multistage dance specific aerobic fitness test (DAFT). J Dance Med Sci. 2003;7(3):80-4.

50. Zoladz JA, Korzeniewski B, Kulinowski P, Zapart-Bukowska J, Majerczak J, Jasiński A. Phosphocreatine recovery overshoot after high intensity exercise in human skeletal muscle is associated with extensive muscle acidification and a significant decrease in phosphorylation potential. J Physiol Sci. 2010;60(5):331-41.