The fallacy of VO_2max

24:4; 15-21, 2009

By Steve Magness

ABSTRACT

This article calls into question the use of VO_2max, both as a physiological parameter for measuring performance capacity and as a variable for planning the training programmes of distance runners and other endurance athletes. Looking first at the rise, starting in the 1920s, of VO_2max as a measurable parameter, the author explains why it has become ingrained in sport and exercise science. This leads to a look into the current research, particularly the Central Governor Model, and a reassessment of the importance of VO_2max and its practical use in training. Citing research that shows basing training based on VO_2max leads to a wide range of individual responses, even among homogenous groups, and that VO_2max does not increase in well trained runners, the author suggests that training at a velocity that corresponds with VO_2max is not the magical training stimulus it is sometimes portrayed to be. This leads to the author's conclusion in which he asks why is so much training focused on a variable that does not change in well trained athletes, barely changes in the moderately trained, levels off after a short period of time, leads to a wide range of adaptations, and does not even correlate well with performance?

AUTHOR

Steve Magness is a student at George Mason University in Washington D.C., USA, where he is currently completing a Masters degree in exercise science with a focus on periodisation in distance running. He is a middle distance runner with a best of 3:43 for 1500m and he coaches a group of national class junior distance runners.

Introduction

In the very first sentence of their comprehensive review of training, MIDGLEY & McNAUGHTON (2006) state: "The maximal oxygen uptake (VO_2max) has been suggested to be the single most important physiological capacity in determining endurance running performance[12]". Based on this notion, training for distance runners and other endurance sports has become fixated on the concept of VO_2max. Training to enhance VO_2max is the subject of numerous review articles and popular coaching material. A whole theory of training has evolved based on the idea of training at the velocity that corresponds with VO_2max, and at certain percentages of VO_2max. Given the emphasis on this particular parameter one would assume that it must be very closely tied with performance and fatigue.

In fact, it is not.

In this paper the limitations of VO_2max will be examined. The discussion will include the legitimacy of the variable itself, why it arose to such prominence, how closely it ties to performance, the efficacy of basing training pace on it and the question of whether athletes should even train to improve it.

How the VO_2max concept developed

The ability to measure oxygen consumption first arose in the early 1920's. It was in 1923, when A.V. Hill and his partner H. Lupton came

up with the idea of an upper limit to oxygen consumption. In an experiment that consisted of Hill running at various speeds around a grass track while measuring VO_2, it was found that he reached a VO_2max of 4.080 L/min at a velocity 243m/min[2]. Despite increases in running velocity, his VO_2 did not increase, leading to the conclusion that there is a maximum limit to oxygen consumption, or in Hill's words:

"In running the oxygen requirement increases continuously as the speed increases attaining enormous values at the highest speeds: the actual oxygen intake, however, reaches a maximum beyond which no effort can drive it… The oxygen intake may attain its maximum and remain constant merely because it cannot go any higher owing to the limitations of the circulatory and respiratory system.[15]"

These findings led to two lasting conclusions. First was that VO_2max is limited by the circulatory and respiratory system. The second was the result of trying to devise a laboratory test for determining VO_2max, in which thirty years later, TAYLOR et al. (quoted in NOAKES, 2008) decided that during a graded exercise test, a VO_2max was obtained when a plateau occurred in VO_2[15]. However, in the authors' original definition, a plateau was not a true plateau, rather it consisted of a VO_2increase of less than 150ml/min from one workload to the next. These findings led to the idea that in order for a true VO_2max to be reached, a plateau of the VO_2 should occur.

Understanding how the VO_2max test came about is important as it impacts the way we currently view and use the parameter. The fact that VO_2max was first measured during exercise by one of the pioneers of exercise science in the 1920's goes a long way in explaining the level of importance ascribed to it.

Whenever a new parameter is discovered or introduced, a large degree of emphasis is put on it in ensuing research. The initial reaction by many scientists is to ascribe a great deal of significance to the newly discovered parameter, as if it will answer all of the questions that they have. It is almost human nature to go through this process of discovery and then exaggeration of the importance of a new finding. Whenever something is new, it is overemphasised until it settles into its rightful place of importance over time. This can be seen in many instances, such as the initial importance put on testing lactate during training.

Due to the very early development of the VO_2max concept, a large amount of early research and study was focused on it, escalating its perceived importance. In addition, from very early on, theories utilising the VO_2max concept were developed and tested, leading to a situation where there was an enormous amount of research and data surrounding them. In essence, a concept too large to be broken down was created and it is almost as if the whole field of exercise science was built upon the VO_2max concept.

Re-evaluating VO_2max

Recently, the legitimacy of VO_2max as a measurement and the acceptance of VO_2max as a practical measurement of cardio-respiratory endurance have been called into question. The contention is that VO_2max is not actually a representative measure of the maximum ability to transport oxygen, but is rather controlled by a central governor. In the Central Governor Model (CGM) developed by NOAKES and others, the CGM predicts that the body regulates exercise to prevent myocardial ischemia. This is accomplished by limiting the blood flow to the periphery, which the brain accomplishes by regulating muscle recruitment[16]. In other words, a central governor acts as a regulator for exercise instead of exercise being limited by some parameter and VO_2max reflects the regulation of muscle recruitment.

There are several theoretical arguments for this model. CGM proponents point to the fact that fatigue is seldom catastrophic, as would be predicted in traditional models. Instead, the body uses various feedback information and past experiences to modulate power out-

put or in the case of running, pace. The idea of pacing being prevalent in endurance events and the fact that a finishing kick, or endspurt, occurs are given as further evidence to support this model[13]. Interestingly, evidence of alterations in pacing strategy and electromyography (EMG), which measures muscle activation, can be seen from the very beginning of performance, such as when athletes are racing in warm versus cool weather, which leads credence to the anticipation model of fatigue[15].

An increase in muscle activation is also seen during the last segments of races, which should not be able to occur if the muscle is "failing" due to fatigue. NOAKES' hypothesis is that at the end of a race, the body's feedback says that it is near completion so that it can push slightly more into its capacity[15]. Evidence for this hypothesis can be seen in a study by TUCKER et al. (2007). They found that in a 20km cycling trial in normoxia versus hyperoxia, the improvement in power output in hyperoxia was proportional to the increase in EMG that also occurred. The authors cited this as evidence that control of muscle activation was one way in which performance was regulated[21].

Another interesting point raised in the CGM debate is the effect hypoxia has on cardiac output. Exercise in hypoxic conditions shows a reduction in peak cardiac output, due to both a decrease in heart rate (HR) and stroke volume (SV)[5]. According to the conventional model, cardiac output, since it is regulated by muscle oxygen demands, should not be reduced. However in the CGM, cardiac output is reduced as a regulatory mechanism and its level is determined by the work done by the muscles[14]. Thus, a reduction in cardiac output in hypoxia is due to a decrease in muscle activation and when supplementary oxygen is taken, cardiac output immediately increases to normal levels[14]. This immediate increase in cardiac output demonstrates that there is a regulatory mechanism in control. Moreover, one has to question why cardiac output is reduced at altitude when oxygen demand by the muscles should be higher.

With regards to VO_2max and how it is tested, NOAKES (2008) has pointed out that in most cases the original requirement of seeing a plateau in VO_2max during an incremental exercise test does not occur[15]. This lack of a plateau is demonstrated in a study of world-class cyclists, where only 47% reached a plateau, prompting the authors of the study to state that the cyclists' limitations might not be oxygen dependent[15]. It is amusing that some authors have commented that motivation may be the reason some athletes do not reach a plateau[18]. This could be a valid statement if the subjects were sedentary, however since the above study was with world-class cyclists, it seems a bit ludicrous to suggest that motivation during a maximum test would be a problem.

In other studies, including one by HAWKINS et al. (quoted in NOAKES, 2008), there have been individual variations in VO_2max levels between the traditional incremental test and a supramaximal test[15]. While in the average of the whole group there were no differences between the tests, the fact that certain individuals showed different VO_2max is interesting and shows that the traditional test does not always give the highest VO_2.

Combining the fact that a plateau does not occur in many subjects and the fact that some individuals reached higher VO_2max values in a supramaximal test than they do in the standard incremental one, the use of the standard incremental VO_2max should be called into question.

Still other studies show that knowing or not knowing when a test or trial will end significantly affects physiological parameters, which lends credence to the aforementioned idea. In a study by BADEN et al. (2005), it was demonstrated that running economy significantly changed, along with rate of perceived effort (RPE), during a submaximal run based on whether the group knew they were running 20 minutes or whether they did not know, even if they ended up running 20 minutes[1]. The VO_2max test is one in which participants

do not have an exact finish distance or time, so it is possible that this degree of uncertainty could affect the physiological parameters measured. The study also points to the importance of feedback and anticipation and that these can affect physiological variables.

Considering the new CGM theory of fatigue, and the fact that the requirement used for reaching VO_2max does not occur in many subjects, the use of VO_2max as a testing parameter must be called into question. In addition, if VO_2max is regulated, then the question arises if it accurately reflects cardiorespiratory endurance. If we accept this, then using VO_2max and percentages of VO_2max for training might not give the training response that we think it does.

Efficacy of basing training paces on VO_2max

With the rise of VO_2max research, training has become based on the parameter in two ways. First, training at the velocities that elicit VO_2max has become the magic training intensity, the speed at which the most improvement can be expected. Second, percentage of VO_2max has become en vogue as a way to quantify training intensity.

In regards to training at VO_2max, this arose because of a review of research that showed that the largest improvements in VO_2max occurred when training at an intensity that corresponded with the parameter, regardless of the duration of the exercise[23]. This finding was subsequently used to demonstrate that training at VO_2max was the best intensity for improving endurance in all groups of people. There are two problems with this conclusion. First, the study's findings are generalised to all groups, even though, as we will talk about later, VO_2max does not improve in well-trained individuals. Second, VO_2max and endurance performance are used almost synonymously, which, as discussed above, is not true. VO_2max may not even measure cardiorespiratory endurance and is certainly not the only factor in endurance performance.

Despite these concerns, training at VO_2max has risen to prominence. It has gone so far, that maximising the time spent at VO_2max has garnered much attention and there are seemingly countless studies and reviews that focus on training at this intensity [12].

Researchers have studied various interval training programmes with the sole goal in seeing how much time at VO_2max each subject spent during the training, which in itself is interesting because it shows the emphasis on the parameter instead of performance. The thought is that time spent at VO_2max is the stimulus needed to improve VO_2max. However, this theory has not been substantiated by research. For instance, a study by BILLAT et al. (1999) found that after four weeks of training using an interval programme designed to elicit time at VO_2max, neither VO_2max nor, more importantly, performance improved[3]. In addition, even in untrained people, the original review by WENGER & BELL (1984) stated that improvements in VO_2max at high intensities were not dependent on the volume of training[23]. Despite these facts, researchers continue to press on with the idea that time spent at VO_2max is the key ingredient for improved endurance, even though no findings back up this theory.

Using percent of VO_2max to quantify intensity is also an accepted practice in research and is used in many training programmes. The problem with this approach is that each individual will have a wide range of adaptation, even if training at the same percentage of VO_2max. This occurs due to differences in the individual's physiology. For instance, lactate threshold can occur at wide range of % VO_2max, even in trained individuals[4]. As an example, if two trained runners both performed at a fixed intensity at 80% VO_2max, one can be below lactate threshold and one above. This would substantially impact the energetics of the workout, as can be seen in a study that showed there was a 40-fold range for increases in lactate levels at 70% VO_2max among individuals[22].

A recent study by SCHARHAG-ROSENBERGER et al. (2009) tested whether exercising at the same %VO_2max resulted in similar metabolic strain[17]. It was found that there is a large individual variance in the lactate response at the fixed intensity, even if groups were matched for similar VO_2max values. This led the authors to conclude that, if the goal is to have similar metabolic strain by subjects in training or research, %VO_2max values should not be used.

In addition to lactate differences, other factors such as the individual's substrate use, fibre type, and other physiological variables will all vary considerably at a fixed percent of VO_2max. A recent study by VOLLAARD et al. (2009) showed that while, on average, improvements were seen in a variety of endurance parameters after six weeks of endurance training, the individuality of response varied widely, with some showing even negative responses to the training, even though the training was at the same 70% VO_2max intensity for all subjects[22]. The study also showed that there was a wide range of adaptation in maximal and submaximal tests, including VO_2parameters, muscle enzyme activity and metabolite levels. An interesting finding is that low responders for an increased VO_2max were not low responders in other parameters. The change in VO_2max did not correlate with the change in performance in a time trial, which is a significant finding demonstrating that perhaps more attention should be paid to changing performance instead of manipulating physiological parameters such as VO_2max.

With this in mind, one has to question the recommendations based on training designed to improve parameters such as VO_2max, with the assumption being that performance will improve because of it, when studies show that change in VO_2max is often not linked with a change in performance. This phenomenon of varied response is not new and can be seen in a wide array of training situations, such as altitude training for example[6]. Moreover, knowing the wide variance in adaptation that can occur when training at a fixed percent of VO_2max, its use has to be called into question. In fact, the VOLLAARD et al. study questioned the use of %VO_2max as a way to standardise intensity and suggested standardisation based on parameters that more directly effect power output.

These findings combined with those by SCHARHAG-ROSENBERGER et al. (2009) suggest that the use of %VO_2max should be eliminated if the goal is to standardise an intensity[17]. One has to really wonder about training programmes that use %VO_2 to prescribe training, as what adaptations will take place are almost a lottery. This does not seem as scientific a way to train as it is portrayed. In practical terms for trained distance runners, it probably makes more sense to standardise paces in relation to their recent race performances or percentages of goal race pace in well-trained runners.

Should we train to improve VO_2max?

As mentioned previously, studies have shown that training at VO_2max elicits the most improvement in VO_2max. This has been used as reasoning for training at VO_2max because, as previously discussed, VO_2max is the traditional measurement for endurance. The logic is that if VO_2max is increased, endurance performance increases. This may not necessarily be the case.

In addition, the question arises if VO_2max actually improves in well-trained runners? In fact, it does not.

Showing the separation of VO_2max and performance, the VOLLAARD et al. study found that the change in VO_2max was not related to the change in time trial performance[22]. Other studies demonstrate improved performances without changes in VO_2max[7]. Still others show that VO_2max can improve without changes in performance, which is seen in a study by SMITH et al. (2003) that showed improvements in VO_2max by 5.0% without an improvement in performance over either 3000m or 5000m[19]. In addition, in looking at long-term changes in performance in

elite athletes, changes in performance occur without subsequent changes in VO_2max.

In highly trained athletes, many studies have shown that VO_2max does not change, even with performance improvements. In one of the only studies done on a large group (33) of elite runners, LEGAZ ARRESE et al. (2005) tracked changes in VO_2max across three years. Performance improved by an average of 1.77% in the men studied, and .69% in the women, with VO_2max remaining essentially unchanged (~76.56 vs. ~76.42 in the men, and ~70.31 vs. ~70.05 in the women)[10]. This points to improved performance in elite runners without changes in VO_2max. Furthermore, it has been shown that among homogenous groups, such as well-trained runners, VO_2max does not correlate well with performance and cannot be used to distinguish which runners are faster[11].

Further evidence can be seen in two case studies on elite runners. In a study on a female Olympic level runner, JONES (1998) showed that while the athlete's 3000m time improved by 46 seconds, the VO_2max decreased from 72 ml/kg/min down to 66 ml/kg/min[8]. Another study by the same author, this one on the current women's marathon world record holder, Paula Radcliffe (GBR) found that while VO_2max varied some based on the time of testing, it was essentially stable at 70 $ml \cdot kg^{-1} \cdot min^{-1}$ from 1992 to 2003[9]. The fact that Radcliffe's VO_2max was essentially stable despite her training volume and intensity increasing substantially is intriguing. Her training increased from a modest 25-30 miles (40-49km) per week (and her VO_2max was already 72 at the time) to 120-160 miles (192-256km) per week. The fact that VO_2max did not change despite this massive increase in volume and intensity points to the short time course of changes in VO_2max.

The rapid change in VO_2max can even be seen in untrained individuals. In a study by SMITH & DONNELL (1984), the authors evaluated the changes in VO_2max over a 36 week training period[20]. VO_2max substantially increased, by 13.6%, but all of the gains were seen in the first 24 weeks of the study with no further increases during the final 12 weeks. Similarly in a study of untrained subjects by DANIELS et al. (1978), VO_2max increased during the first four weeks of training but did not increase after that, even though there was a further increase of training and continued improvements in performance[7]. Given the evidence that VO_2max does not change in elite runners and does not correlate with performance, training focused on improving VO_2max does not seem like a logical approach for coaches of well-trained runners.

Conclusion

VOLLAARD et al. may have put it best when they came to the conclusion that "Moreover, we demonstrate that VO_2max and aerobic performance associate with distinct and separate physiological and biochemical endpoints, suggesting that proposed models for the determinants of endurance performance may need to be revisited[22]". Their recognition that aerobic performance and VO_2max are not direct equals or even well linked is a step in the right direction and needs to be acknowledged to a much greater degree.

Combining these findings with NOAKES' CGM creates a situation where VO_2max may not be measuring what we think it is. Adding the facts that using $\%VO_2$ to classify training results in a wide range of adaptations and that changes in VO_2max do not occur in trained athletes, one has to question the reason for basing entire training programmes on VO_2max.

The bottom line question for scientists and coaches is: why is so much of the training for endurance focused on a variable that does not change in well trained athletes, barely changes in the moderately trained, levels off after a short period of time, and does not even correlate well with performance?

Please send all correspondence to:
Steve Magness
sjm1368@yahoo.com

REFERENCES

1. BADEN, D. A.; McLEAN, T. L.; TUCKER, R.; NOAKES, T. D. & ST CLAIR GIBSON, A. (2005). Effect of anticipation during unknown or unexpected exercise duration on ratings of perceived exertion, affect, and physiological function. British Journal of Sports Medince, 39(1), 742–746.
2. BASSETT, D. R. & HOWLEY, E. T. (2000). Limiting factors for maximum oxygen uptake and determinants of endurance performance. Medicine and Science in Sports and Exercise, 32, 70–84.
3. BILLAT, V. L.; FLETCHET, B.; PETIT, B.; MURIAUX, G. & KORALSZTEIN, J. P. (1999). Interval training at VO2max effects on aerobic performance and overtraining markers. Medicine and Science in Sports and Exercise, 31, 156–163.
4. BROOKS, G. A.; FAHEY, T. D. & BALDWIN, K. (2004). Exercise Physiology: Human bioenergetics and its application. McGraw-Hill.
5. CALBET, J. A.; BOUSHEL, R.; RADEGRAN, G.; SONDERGAARD, H.; WAGNER, P. D. & SALTIN, B. (2003). Determinants of maximal oxygen uptake in severe acute hypoxia. Am J Physiol Regul Integr Comp Physiol, 284(2), 291–303.
6. CHAPMAN, R.; STRAY-GUNDERSON, J. & LEVINE, B. D. (1998). Individual variation in response to altitude training. Journal of Applied Physiology, 85(4), 1448–1456.
7. DANIELS, J. T.; YARBROUGH, R. A. & FOSTER, C. (1978). Changes in VO_2max and running performance with training. European Journal of Applied Physiology, 39(4), 249–254.
8. JONES, A.M. (1998). A five year physiological case study of an Olympic runner. British Journal of Sports Medicine 32: 39–43.
9. JONES, A.M (2006). The physiology of the world record holder for the women's marathon. International Journal of Sports Science and Coaching 1,101–116.
10. LEGAZ ARRESE, A.; SERRANO OSTÁRIZ, E.; JCASAJÚS MALLÉN, J. A. & MUNGUÍA IZQUIERDO, D. (2005). The changes in running performance and maximal oxygen uptake after long-term training in elite athletes. Journal of Sports Medicine and Physical Fitness, 45(4), 435–40.
11. LEGAZ ARRESE, A.; MUNGUÍA IZQUIERDO, D.; NUVIALA NUVIALA, A.; SERVETO-GALINDO, O.; MOLINER URDIALES, D. & REVERTER MASIA, J. (2007). Average VO_2max as a function of running performances on different distances. Science & Sports, 22(1), 43–49.
12. MIDGLEY, A.W.; McNAUGHTON, L. R. & WILKINSON, M. (2006). Is there an optimal training intensity for enhancing maximal oxygen uptake of distance runners? Empirical research findings, current opinions, physiological rationale and practical recommendations. Sports Medicine, 36(2), 117–132.
13. NOAKES, T. D. (2003). Commentary to accompany training and bioenergetic characteristics in elite male and female Kenyan runners. Medicine and Science in Sports and Exercise, 35(2), 305–306.
14. NOAKES, T. D.; CALBET, J. A.; BOUSHEL, R.; SONDERGAARD, H.; RADEGRAN, G.; WAGNER, P. D. & SALTIN, B. (2004). Central regulation of skeletal muscle recruitment explains the reduced maximal cardiac output during exercise in hypoxia. Am J Physiol Regul Integr Comp Physiol, 287(4) R996-999. author reply R999–1002.
15. NOAKES, T. D. (2008). How did A.V. Hill understand the VO_2max and the "plateau phenomenon"? Still no clarity? British Journal of Sports Medicine, 42(7), 574–580.
16. NOAKES, T. D. & MARINO, F. E. (2009). Point: counterpoint: maximal oxygen uptake is/is not limited by a central nervous system governor. Journal of Applied Physiology, 106, 338–339.
17. SCHARHAG-ROSENBERGER, F.; MEYER, T.; GABLER, N.; FAUDE, O. & KINDERMANN, W. (2009). Journal of Science and Medicine in Sport, in press.
18. SHEPARD, R. J. (2009). Is it time to retire the 'central governor'. Sports Medicine, 39(9), 709–721.
19. SMITH T. P.; COOMBES, J. S. & GERAGHTY, D. P. (2003). Optimising high-intensity treadmill training using the running speed at maximal O(2) uptake and the time for which this can be maintained. European Journal of Applied Physiology, 89(3-4), 337–343.
20. SMITH, D. A. & O'DONNELl, T. V. (1984). The time course during 46 weeks' endurance training of changes in Vo2max and anaerobic threshold as determined with a new computerized method. Clinical Science, 67(2), 229–236.
21. TUCKER, R.; KAYSER, B.; RAE, E.; RAUNCH, L.; BOSCH, A. & NOAKES, T. D. (2007). Hyperoxia improves 20 km cycling time trial performance by increasing muscle activation levels while perceived exertion stays the same. European Journal of Applied Physiology, 101(6), 771–781.
22. VOLLAARD, N. B. J.; CONSTANTIN-TEODOSIU, D.; FREDRIKSSON, K.; ROOYACKERS, O.; JANSSON, E.; GREENHAFF, P. L.; TIMMONS, J. A. & SUNDBERG, C. J. (2009). Systematic analysis of adaptations in aerobic capacity and submaximal energy metabolism provides a unique insight into determinants of human aerobic performance. Journal of Applied Physiology, 106, 1479–1486.
23. WENGER, H. A. & BELL, G. J. (1986). The interactions of intensity, frequency and duration of exercise training in altering cardiorespiratory fitness. Sports Medicine, 3(5), 346–356.