Accuracy of $\dot{V}O_{2max}$ and anaerobic threshold determination*

Paulo Cesar N. Granja Filho, Fernando A.M.S. Pompeu and Alair Pedro Ribeiro de Souza e Silva

ABSTRACT

Studies on the accuracy of the gas exchange and ventilation parameters during exertion involving the Brazilian population are scarce in literature. **Objective:** To determine the reliability of the maximal oxygen intake ($\dot{V}O_{2max}$) and anaerobic threshold (AnT), as well as the objectivity of the AnT determination in Brazilian healthy youth adults. **Methods:** Two tests of maximal exertion were applied. Two independent observers applied the visual inspection method for the AnT determination. The data were compared by means of the regression analysis, intraclass correlation coefficient (ICC), two-way ANOVA and the paired t-test for an $\alpha \leq 0.05$. The intra-observer and intra-subject variabilities were determined by means of the typical error (s) and variation coefficient (VC). **Results:** No significant differences between tests for the $\dot{V}O_{2max}$ and AnT determination as well as between observers in AnT determination were observed. The new $\dot{V}O_{2max}$ presented ICC = 0.97, s = ± 0.14 L • min^{-1}, and VC = ± 5.5%. The AnT presented ICC = 0.90, s = ± 0.14 L • min^{-1}, and VC = ± 9.2%. The inter-observers AnT determination presented ICC = 0.95, s = 0.1 L • min^{-1} and VC = 5.6%. **Conclusion:** The $\dot{V}O_{2max}$ and the AnT are reliable measurements, and the AnT determination was demonstrated to be an objective method in the population studied.

Key words: Reliability. Objectivity. Maximal oxygen intake. Anaerobic threshold and ventilatory threshold.

INTRODUCTION

The maximum oxygen intake ($\dot{V}O_{2max}$) and the aerobic threshold (AnT) are parameters used in the evaluation of the maximum cardiorespiratory function and functional reserve. The $\dot{V}O_{2max}$ has aided on the selection of candidates for heart transplantation (< 14 mL • kg^{-1} • min^{-1}), once it presents high predictive value with regard to the mortality rate of these patients[1,2].

The athletic performance in long duration events presents good association (r = 0.78) with $\dot{V}O_{2max}$[3]. However, this phenomenon was not observed in homogeneous groups of high-level athletes[4,5]. Stronger correlations may be found between AnT and the performance in long distance events (r = 0.94 to 0.98)[3].

In order to measure these parameters in the clinical and sportive scope, automatic ergospirometers with good temporal resolution are employed. However, the instruments available in most national laboratories are less sophisticated due to their lower cost. These equipments have not been evaluated with regard to the quality of measurements. In this case, Brazilian researchers are limited to extrapolations based on foreign studies performed with sophisticated equipments. Therefore, the objective of the present investigation was to analyze the $\dot{V}O_{2max}$ and AnT reliability as well as the objectivity of AnT, when equipments widely used in the national territory are employed.

METHODS

Fourteen non-smoker healthy young adults involved in non-competitive physical activities were volunteers in this study: 24 ± 4 years of age, 66.2 ± 13.7 kg and 169.6 ± 10.2 cm. This group was composed of five women (28 ± 6 years of age; 57.2 ± 4.1 kg and 160.2 ± 7.1 cm) and nine men (22 ± 2 years of age, 71.1 ± 14.8 kg and 174.8 ± 7.4 cm). Before tests, all volunteers filled a Cleared and Free Consent Term. The experimental protocol of this study was previously approved by the Ethics Committee for studies involving human beings of the HSE (CEP. 000.021/99). In the day before the examination, the participants were recommended not to perform exhausting physical activities. They were also recommended to avoid caffeine and food three hours before exertion test.

At the first visit to laboratory, the volunteers performed a familiarization test. One week later, the subjects returned to laboratory for the performance of the experimental tests, which were separated by the minimum interval of one and maximum interval of two weeks.

A assigned and continuous protocol[6] was used, which was composed of initial rest for six minutes sitting on the cycle ergometer (*Monark*®, Brazil) followed by a four-minute warm-up exercise with no load and later by the assigned phase with duration of 8-12 minutes. In relation to the increments, the maximum load was calculated (W_{max})[7] and divided into ten parts for the determination of the one-minute stages. The rhythm ranged according to the individual from 60 to 96 rpm and from 15 to 35 W • min^{-1}.

Gas exchange and ventilation variables were integrated each 20 seconds based on measurements collected through a metabolic analyzer (*Aerosport*® TEEM 100, USA) and average-flow pneumotachograph (*Hans Rudolph*®, USA). The heart rate was controlled by means of a cardiotachometer (*Polar Sport Tester*®, Finland) each five seconds.

The equipments calibration procedures were performed prior to the performance of each test. The ergospirometer was calibrated by means of a standard gas mixture (AGA®, Brazil) balanced with nitrogen, containing 17.01% of oxygen and 5.00% of carbon dioxide. The outflow was calibrated through a three-liter syringe (*Hans Rudolph*®, USA) and the cycle ergometer by means of a 3-kg ballast.

The tests, always performed by the same observer were considered as satisfactory when at least one of the following criteria[8] for maximum exertion was reached: $\dot{V}O_2$ plateau (increase ≤ 150 mL • min^{-1} or 2 mL • kg^{-1} • min^{-1}), $HR_{max} \geq 180$ bpm, Borg ≥ 18 and RER ≥ 1.1. There was no verbal or other external nature encouragement that could change the exertion continuation time.

The AnT determination was performed through the visual inspection method[6,9-11]. To do so, the modified v-slope graphics (V-s mod), the ventilatory equivalents ($\dot{V}E/\dot{V}O_2$ x time and $\dot{V}E/\dot{V}CO_2$ x time), the minute ventilation (VE x time) and the gas exchange ratio (RER x time) were analyzed. Two independent observers de-

* Departamento de Biociências da Atividade Física – EEFD/UFRJ.

Received in 17/12/04. 2nd version received in 18/1/05. Approved in 12/5/05.

Correspondence to: Departamento de Biociências da Atividade Física – EEFD/UFRJ, Av. Brigadeiro Trompowisk, s/n, Cidade Universitária, Ilha do Fundão – 29141-590 – Rio de Janeiro, RJ. Fax: (21) 2562-6801. E-mail: pompeu_fernando@hotmail.com

termined the AnT for each subject with the analysis of four graphics all in all (AnT_{app}). Such observers presented a minimum experience of two years in ergospirometry training. The average of values between both observers was used as combined method[12] (AnT_{comb}). The threshold values were expressed as $\dot{V}O_2$abs (L • min^{-1}), $\dot{V}O_2$rel (mL • kg^{-1} • min^{-1}) and physical power (watts).

The statistical treatment was conducted through the *SPSS*® software for *Windows* and *Excel*® spreadsheet[13]. For the analysis of the $\dot{V}O_{2max}$ and AnT_{comb} reliability, the differences between averages were evaluated through the paired t-test. Two two-factor ANOVAs (4 x 4) were used in order to compare graphics for the AnT determination in both tests, or by both observers. The *post hoc* HSD-Tukey test was applied whenever necessary. The association degree between tests and observers was measured through the intraclass correlation coefficient (ICC). The linear analysis regression was used to establish the load (watts) corresponding to the AnT_{comb} (L • min^{-1}) for each subject and yet, to relate $\dot{V}O_{2max}$ and AnT_{comb} between both tests. The intra-subject and intra-observer variations were measured through the typical error (s) and through the variation coefficient (VC)[14]. The significance level adopted was lower than or equal to 0.05.

RESULTS

The difference between $\dot{V}O_{2max}$ in both moments ($\dot{V}O_{2max}$ abs: test A = 3.03 ± 0,81 L • min^{-1} and test B = 3.01 ± 0.85 L • min^{-1}; $\dot{V}O_{2max}$rel: test A = 46.1 ± 9.5 mL • kg^{-1} • min^{-1} and test B = 45.5 ± 9.7 mL • kg^{-1} • min^{-1}) was not significant. The intra-subjects variation for $\dot{V}O_{2max}$ was of 0.14 L • min^{-1} (VC = 5.5%) or 2.0 mL • kg^{-1} • min^{-1} (VC = 5.4%). A strong association degree was observed between both tests for the measures of $\dot{V}O_{2max}$ ($\dot{V}O_{2max}$abs: ICC = 0.97, $\dot{V}O_{2max}$rel: ICC = 0.95).

The differences in the AnT_{comb} between tests, expressed as $\dot{V}O_2$abs (test A = 1.55 ± 0.49 L • min^{-1} and test B = 1.51 ± 0.42 L • min^{-1}), as $\dot{V}O_2$rel (test A = 23.5 ± 5.8 mL • kg^{-1} • min^{-1} and test B = 22.9 ± 5.5 mL • kg^{-1} • min^{-1}) or as load (test A = 112 ± 35 W and test B = 111 ± 30 W), were not significant. Not significant differences were also observed for the AnT determination through each parameter employed in the visual inspection (table 1). The intra-subject variation for AnT_{comb} was of 0.14 L • min^{-1} (VC = 9.2%) or 2.0 mL • kg^{-1} • min^{-1} (VC = 8.5%). A strong association was observed between tests for the AnT_{comb} expressed as $\dot{V}O_2$abs (ICC = 0.90), as $\dot{V}O_2$rel (ICC = 0.87) and as load (ICC = 0.80). The association degree between tests and the intra-observer variation for visual inspection methods are presented in table 2.

TABLE 1
Two-factor ANOVA for comparison between methods in tests A and B for observers

Variable	Tests	Observers	Methods (average ± SD)				
			mod.V-slope	Ve/Vo_2 x time	Ve x time	RER x time	AnT *observer*
AnT ($\dot{V}O_2$, L • min^{-1})	A	A	1.5 ± 0.5	1.6 ± 0.5	1.6 ± 0.5	1.6 ± 0.5	1.6 ± 0.5
	A	B	1.5 ± 0.6	1.5 ± 0.5	1.3 ± 0.3	1.5 ± 0.6	1.4 ± 0.4
	B	A	1.5 ± 0.5	1.6 ± 0.4	1.6 ± 0.4	1.4 ± 0.4	1.5 ± 0.4
	B	B	1.5 ± 0.6	1.4 ± 0.4	1.3 ± 0.3	1.6 ± 0.5	1.4 ± 0.4
AnT ($\dot{V}O_2$, mL • kg^{-1} • min^{-1})	A	A	23.3 ± 5.3	25.1 ± 7.0	24.3 ± 6.0	24.4 ± 5.3	24.5 ± 5.5
	A	B	23.4 ± 8.1	23.0 ± 6.7	21.0 ± 4.4	22.9 ± 8.2	22.5 ± 6.2
	B	A	22.9 ± 6.3	24.7 ± 5.5	24.6 ± 6.0	22.0 ± 4.5	23.9 ± 5.3
	B	B	24.0 ± 8.3	21.9 ± 6.3	20.8 ± 5.5	24.7 ± 8.0	22.2 ± 6.3
AnT (watts)	A	A	127 ± 54	135 ± 50	138 ± 47	138 ± 59	118 ± 41
	A	B	122 ± 60	114 ± 48	112 ± 33	110 ± 46	104 ± 31
	B	A	119 ± 50	136 ± 41	136 ± 36	122 ± 36	117 ± 31
	B	B	119 ± 52	108 ± 45	108 ± 37	131 ± 52	106 ± 33

* difference between methods ($\alpha \leq 0.05$), ** difference between tests ($\alpha \leq 0.05$), *** interaction ($\alpha \leq 0.05$).

TABLE 2
Intra-observer reliability and variability of AnT (expressed as $\dot{V}O_2$ abs, $\dot{V}O_2$ rel, and watts)

Variable	Methods	Observers					
		s		*VC*		*ICC*	
		A	B	A	B	A	B
$\dot{V}O_2$abs (L • min^{-1})	mod V-s	0.17	0.25	13.00%	21.00%	0.89	0.83
	Ve/Vo_2 x time	0.16	0.16	11.40%	11.00%	0.88	0.89
	Ve x time	0.18	0.17	12.10%	13.00%	0.87	0.74
	RER x time	0.22	0.30	12.20%	18.70%	0.88	0.73
	AnT observer	0.11	0.17	7.50%	12.30%	0.94	0.86
		Average ± SD: 0.17 ± 0.03	Average ± SD: 0.21 ± 0.06	Average ± SD: 11.2 ± 2.1%	Average ± SD: 15.2 ± 4.4%	Average ± SD: 0.89 ± 0.02	Average ± SD: 0.81 ± 0.07
$\dot{V}O_2$rel (mL • kg^{-1} • min^{-1})	mod V-s	2.47	3.55	12.70%	17.70%	0.82	0.81
	Ve/Vo_2 x time	2.47	2.13	10.80%	10.50%	0.84	0.89
	Ve x time	3.65	2.52	16.10%	13.00%	0.63	0.75
	RER x time	2.61	4.46	11.50%	17.80%	0.81	0.70
	AnT observer	1.47	2.56	6.70%	11.80%	0.92	0.83
		Average ± SD: 2.53 ± 0.77	Average ± SD: 3.04 ± 0.9	Average ± SD: 11.5 ± 0.03%	Average ± SD: 14.1 ± 0.03%	Average ± SD: 0.8 ± 0.1	Average ± SD: 0.8 ± 0.07
Watts	mod V-s	34	27	29.40%	29.60%	0.57	0.76
	Ve/Vo_2 x time	11	18	9.20%	13.20%	0.93	0.84
	Ve x time	20	17	14.90%	16.00%	0.75	0.75
	RER x time	28	24	15.90%	18.90%	0.77	0.77
	AnT observer	14	16	9.40%	16.30%	0.84	0.74
		Average ± SD: 21 ± 9.5	Average ± SD: 20 ± 4.8	Average ± SD: 15.7 ± 0.08%	Average ± SD: 18.8 ± 0.06%	Average ± SD: 0.77 ± 0.13	Average ± SD: 0.77 ± 0.03

The difference between observers for methods alone or together (table 1) was not significant. A high association degree of AnT was observed between observers in test A expressed as $\dot{V}O_2$abs (ICC = 0.92), as $\dot{V}O_2$rel (ICC = 0.93) and as load (ICC = 0.85). Similar results were also observed in test B when AnT_{obs} was expressed as $\dot{V}O_2$abs (ICC = 0.93), as $\dot{V}O_2$rel (ICC = 0.91) and as load (ICC = 0.90).

DISCUSSION

The combined method for the AnT detection proposed by Gaskill *et al.*(12) was based on the averages of intensities determined by independent observers. In order to establish the AnT intensity, these observers employed the visual inspection of three graphics (simplified v-slope, ventilatory equivalents and excess of exhaled CO_2). In this study(12), a strong association (r^2 = 0.93, n = 54) between test and retest was demonstrated.

Caiozzo *et al.*(10) found excellent correlation between tests for the ventilatory equivalents method (r = 0.93). However, this parameter presented lower reproducibility in the tests of Cohen-Solal *et al.*(16) with cardiac patients (r = 0.83). Davis *et al.*(17) reported moderate correlation (r = 0.74) between tests for AnT using the visual inspection of three graphics (VE x time, FeO_2 x time and $\dot{V}CO_2$ x time). In the last study(17), as in the current one, the determination of $\dot{V}O_2$ in AnT was less reliable than when performed through the respiratory exchange ratio method (RER).

Table 2 presents the anaerobic threshold per observer (AnT_{obs}) using all graphics. One observes that the concurrent employment of all graphics produces lower intra-observer variation when compared to the employment of graphics alone.

Studies on the objectivity of AnT presented conflicting results. Posner *et al.*(17) found ICC = 0.94 between the best observers, Gladden *et al.*(18) reported ICC = 0.70 and Shimizu *et al.*(19), ICC = 0.85.

The time interval between tests and the AnT detection method may present relevant effects in methodological studies. Hopkins(20) suggested, based on a meta-analysis, that the optimum interval between tests should be of 2.5 days. This author believes that this period presents no fatigue residue and therefore no alterations in the second test. However, the interval from 7 to 14 days may be operationally more worthwhile, as, according to the experience from our laboratory, this interval improves the adherence to these tests. In intervals of up to two weeks, no significant changes in the performance of individuals who do not participate in training programs are expected.

The influence of the biological variation on the AnT and $\dot{V}O_{2max}$ reliability is not fully known. Tests as that performed by Katch *et al.*(21) tried to measure this variation through multiple tests along the period of four weeks, five days a week in well-trained athletes (n = 5). These authors concluded that the biological variation is equivalent to 90% of all fluctuation of the $\dot{V}O_{2max}$ results. Interestingly, figure 1 shows that individuals with higher $\dot{V}O_{2max}$ values are the nearest to the identity line. We believe that the biological variation is lower among the most capable individuals. In addition, one observes that the magnitude of the peripheral and central chronic adaptations observed after training are inversely proportional to the initial conditioning level. Skinner *et al.*(22) studied 614 individuals and demonstrated that those with lower $\dot{V}O_{2max}$ presented the highest adaptations (21.9%). In another study(15), the authors demonstrated that tests conducted with well-trained athletes are more accurate than those conducted with sedentary individuals.

The plateau criterion for the $\dot{V}O_{2max}$ characterization was firstly proposed by Taylor *et al.*(23) in the middle of the XXth century and since then it was considered as the main reference of maximum exertion. However, in the present study, this criterion was observed in only 41.6% of the tests. A possible explanation for the absence of this phenomenon may be the time employed in the sampling of gases. Myers *et al.*(24) found large variation on the angular coefficient of the relation between $\dot{V}O_2$ and exertion time when administered several intervals for the collection of gas samples. In this case, the plateau, defined as an inclination smaller than or equal to zero in the $\dot{V}O_2$, was more affected in shorter intervals. These data suggest that the variation in the $\dot{V}O_2$ x time relation depends on the gas samples collection interval.

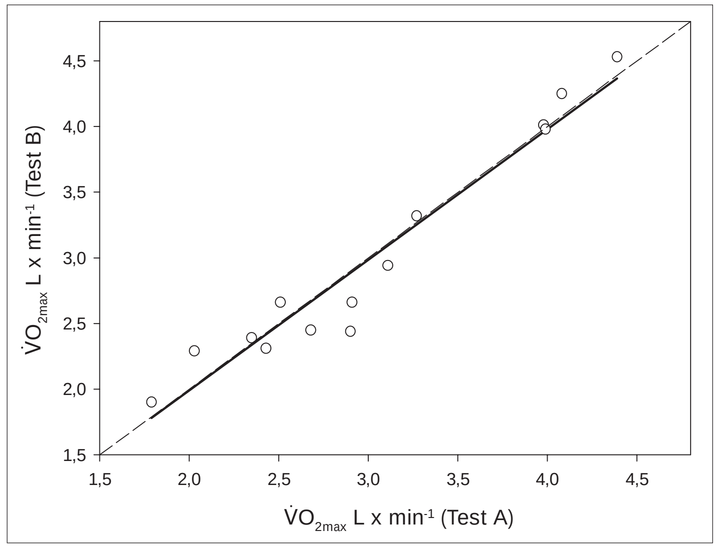

Fig. 1 *– Relation between tests for $\dot{V}O_{2max}$ ($\dot{V}O_2$ abs)*

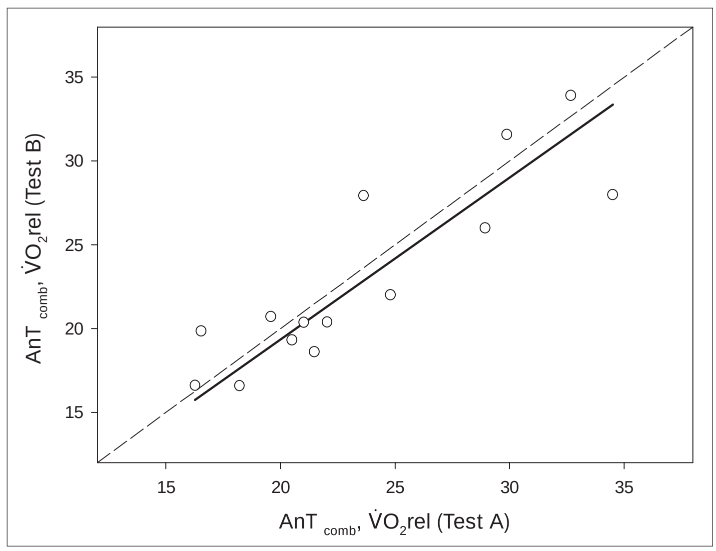

Fig. 2 *– Relation between tests for AnT_{comb} ($\dot{V}O_2$ rel)*

Another possible source of error in the present study may be due to the mechanical-braking cycle ergometer (friction). This type of equipment requires periodic adjustments in the resistance(25). This occurs due to the heating of the nylon belt that modifies the friction and hence the resistance(25). The intra-observer error in the reading of the load calibrated scale in kp may also present some influence on the AnT_{comb} variation. Wilmore *et al.*(26) investigated by electronic means the lack of calibration of different cycle ergometers and the effect of this error on the $\dot{V}O_2$ reliability in several exertion intensities. Those authors observed a variation of approximately 10% in the load of the mechanical-braking cycle ergometer; however, this variation was not sufficient to change the good association (r = 0.88) in the $\dot{V}O_2$ measurement between two moments at load of 98 W. In the investigation of Wilmore *et al.*(26), the decreasing tendency of correlations with the increase on the work load was also observed. In the present study, the averages of loads corresponding to AnT_{comb} were found near to those presented in the study mentioned above[26] (ICC = 0.80). The reduction on the correlation for $\dot{V}O_2$ with the increase on the work load may be explained by the lack of calibration of the ergometer and/or ergospirometer.

When AnT_{comb} was expressed by the work load, a higher variation between both moments was observed (VC = 12.1%) if compared to that expressed by the $\dot{V}O_2$ ($L \cdot min^{-1}$: VC = 9.2%; and as $mL \cdot kg^{-1} \cdot min^{-1}$: VC = 8.4%). In several studies, VC values from 1.9 to 8.4% for the $\dot{V}O_{2max}$ were observed[21,23,27-29]. Gerrard *et al.*[29] included the diurnal variation in the treatment of their data. Those authors performed six tests for each subject (3 days x 2 periods) and could observe VC of 8.4% for the $\dot{V}O_{2max}$ and of 12% for the AnT. Taylor *et al.*[23] obtained VC = 2.4% using discontinuous protocol and measuring $\dot{V}O_2$ in the last minute of the test under well-controlled conditions. The subjects' motivation may be an important aspect for the reliability of the $\dot{V}O_{2max}$ measure[30]. However, the AnT theoretically should not be affected by motivation, once it occurs within submaximal effort levels.

Due to the Brazilian cultural and socioeconomic conditions, it is possible that the lack of experience with cyclism may have influenced negatively the results obtained. Cyclism produces distinct effects on $\dot{V}O_{2max}$ in sedentary individuals (generally 5 to 20% lower than that observed in running) and in cyclists (generally 10% higher than that observed in running). In case the subjects had been submitted to treadmill test, a higher reliability would have possibly been obtained.

Cautions in the pre-tests may also have influenced the results. However, the subjects were informed about the importance of non-engaging in physical training programs, once significant changes on AnT are expected (expressed as $L \cdot min^{-1}$, but not as $\%\dot{V}O_{2max}$) in normal men from the third week of training at 80% of the $\dot{V}O_{2max}$ on, during 30 minutes, four days a week[31].

Therefore, one concludes that the $\dot{V}O_{2max}$ and the AnT present small intra-subject, intra-observer and inter-observer variation (in the second case), being precise parameters when the equipments used are those most frequently adopted in national laboratories.

ACKNOWLEDGMENTS

The authors would like to thank the financial support and the priceless academic contribution from AACEA-HSE in the person of Dr. Aluysio S. Aderaldo Jr. They also thank the contribution of academicians Jan B. Bartholomeu and Paulo André da Silva in the data collection phase. Finally, they would like to thank the valuable text revisions by PhD Profs. Martha M. Sorenson and Verônica S. Pinto (Bioqmed/UFRJ).

All the authors declared there is not any potential conflict of interests regarding this article.

REFERENCES

1. Costanzo MR, Augustine S, Bourge R, Bristow M, O'Connell JB, Driscoll D, et al. Selection and treatment of candidates for heart transplantation. Circulation 1995;92:3593-612.
2. Mancini DM, Einsen H, Kussmaul W, Mull R, Edmunds Jr LH, Wilson JR. Value of peak exercise oxygen consumption for optimal timing of cardiac transplantation in ambulatory patients with heart failure. Circulation 1991;83:778-86.
3. Sjödin B, Svedenhag J. Applied physiology of marathon running. Sports Med 1985;2:83-99.
4. Coetzer P, Noakes TD, Sanders B, Lambert MI, Bosch NA, Wiggins T, et al. Superior fatigue resistance of elite black South African distance runners. J Appl Physiol 1993;75:1822-7.
5. Weston AR, Mbambo Z, Myburgh KH. Running economy of African and Caucasian distance runners. Med Sci Sports Exerc 2000;32:1130-4.
6. Wasserman K, Whipp BJ, Koyal SN, Beaver WL. Anaerobic threshold and respiratory gas exchange during exercise. J Appl Physiol 1973;35:236-43.
7. Hsi WL, Lan C, Lai JS. Normal standards for cardiopulmonary responses to exercise using a cycle ergometer test. J Formos Med Assoc 1998;97:315-22.
8. Howley ET, Basset Jr DR, Welch HG. Criteria for maximal oxygen uptake: review and commentary. Med Sci Sports Exerc 27;1292-301.
9. Sue DY, Wasserman K, Morrica RB, Casaburi R. Metabolic acidosis during exercise in patients with chronic obstructive pulmonary disease: use of V-slope method for anaerobic threshold determination. Chest 1988;94:931-8.
10. Caiozzo VJ, Davis JA, Ellis JF, Azus JL, Vandagriff R, Prietto CA, et al. A comparison of gas exchange indices used to detect the anaerobic threshold. J Appl Physiol 1982;53:1184-9.
11. Wasserman K, Mcilroy M. Detecting the threshold of anaerobic metabolism in cardiac patients during exercise. Am J Cardiol 1964;14.
12. Gaskill S, Ruby BC, Walker AJ, Sanchez AO, Serfass RC, Leon AS. Validity and reliability of combining three methods to determine ventilatory threshold. Med Sci Sports Exerc 2001;33:1841-8.
13. A New View of Statistics [homepage on the internet]. Hopkins WG. Reliability from consecutive pairs of trials [updated 2003; cited 2003 Set 06]. Available from: http:\\sportsci.org/resource/stats/xrely.xls
14. Hopkins WG. Measures of reliability in sports medicine and science. Sports Med 2000;30:1-15.
15. Cohen-Solal A, Benessiano J, Himbert D, Paillole C, Gourgon R. Ventilatory threshold during exercise in patients with mild to moderate chronic heart failure: determination, relation with lactate threshold and reproducibility. Int J Cardiol 1991; 30:321-7.
16. Davis JA, Vodak P, Wilmore JH, Vodak J, Kurtz P. Anaerobic threshold and maximal aerobic power for three modes of exercise. J Appl Physiol 1976;41:544-50.
17. Posner JD, Gorman KM, Klein HS, Cline CJ. Ventilatory threshold: measurement and variation with age. J Appl Physiol 1987;63:1519-25.
18. Gladden LB, Yates JW, Stremel W, Stanford BA. Gas exchange and lactate anaerobic thresholds: inter- and intraevaluator agreement. J Appl Physiol 1985;58:2082-9.
19. Shimizu M, Myers J, Buchanan N, Walsh D, Kraemer M, McAuley P, et al. The ventilatory threshold: method, protocol, and evaluator agreement. Am Heart J 1991;122:509-16.
20. Hopkins WG, Schabort EJ, Hawley JA. Reliability of power in physical performance tests. Sports Med 2001;31:211-34.
21. Katch VL, Sady SS, Freedson P. Biological variability in maximum aerobic power. Med Sci Sports Exerc 1982;14:21-5.
22. Skinner JS, Wilmore KM, Krasnoff JB, Jakolski A, Gagnon J, Province MA, et al. Adaptation to standardized training program and changes in fitness in large, heterogeneous population: the heritage family study. Med Sci Sports Exerc 2000; 32:157-61.
23. Taylor HL, Buskirk E, Henschel A. Maximal oxygen intake as objective measure of cardiorespiratory performance. J Appl Physiol 1955;8:73-80.
24. Myers J, Walsh D, Sullivan M, Froelicher V. Effect of sampling on variability and plateau in oxygen uptake. J Appl Physiol 1990;68:404-10.
25. Paton CD, Hopkins WG. Tests of cycling performance. Sports Med 2001; 31:489-96.
26. Wilmore JH, Constable SH, Stanforth PR, Buono MJ, Tsao YW, Roby JR, et al. Mechanical and physiological calibration of four cycle ergometers. Med Sci Sports Exerc 1982;14:322-5.
27. Kuipers H, Verstappen FT, Keizer HA, Geurten P, Van KG. Variability of aerobic performance in the laboratory and its physiologic correlates. Int J Sports Med 1985;6:197-201.
28. Figueroa-Colon R, Hunter GR, Mayo MS, Aldridge RA, Goran ML, Weinsier RL. Reliability of treadmill measures and criteria to determine $\dot{V}O_{2max}$ in prepubertal girls. Med Sci Sports Exerc 2000;32:865-9.
29. Gerrard CS, Emmons C. The reproducibility of respiratory responses to maximum exercise. Respiration 1986;49:94-100.
30. Shephard R. Maximal oxygen intake. Exercise in sports. Australia: Blackwell Scientific Publications, 1992;198.
31. Ready AE, Quinney HA. Alterations in anaerobic threshold as the result of endurance training and detraining. Med Sci Sports Exerc 1982;14:292-6.