Original Article

Validation of four indirect VO₂max laboratory prediction tests in the case of soccer players

RATKO PERIC¹, ZORAN NIKOLOVSKI²

¹Institute for Sport science, University of Graz, AUSTRIA

²Sport Science Department, Aspire Academy, Doha, QATAR

Published online: June 30, 2017

(Accepted for publication June 11, 2017)

DOI:10.7752/jpes.2017.02092

Abstract:

The aim of this study was to validate four commonly used tests (Astrand-Ryhming, Bruce, Balke and PWC₁₅₀) for indirect maximum oxygen uptake (VO₂max) prediction and to assess their usage in the case of soccer players. Sixteen male subjects (age = 27.8 ± 3.8 y, height = 182.8 ± 4.7 cm and body mass = 81.5 ± 8.5 kg) participated in this study. Four observed tests were compared with direct VO₂max assessment which was performed as breath-by-breath gas analysis during graded treadmill protocol until exhaustion. The Astrand-Ryhming and PWC₁₅₀ tests were performed on a stationary bike while the Bruce and Balke tests were conducted on a motorized treadmill. The methods of statistical analysis and comparison used were Bland-Altman plot analysis, regression analysis, the coefficient of variation and one way ANOVA. Our results revealed no agreement among measurement techniques and mean differences in all four protocols. The relationship was modest but statistically significant in the case of Bruce protocol ($R^2 = 32\%$, $p \leq 0.05$). We conclude moderate reliability to be sufficiently satisfactory for the observed population and can be used under given circumstances.

Key words: Indirect, VO₂max, CPET, Ergometry, Soccer.

Introduction

Cardiopulmonary exercise testing with direct measurement of VO₂ (CPET) is widely considered as a golden standard for maximum oxygen uptake (VO₂max) and fitness level assessment. However, this method is expensive, requires specialized staff and highly demanding preconditions need to be met (American Thoracic Society, 2003). Regrettably, not all countries meet these requirements, making CPET unavailable. Under such circumstances, especially in studies of large population, indirect VO₂max prediction needs to be implemented. For indirect VO₂max prediction tests, the linear relation between heart rate (HR) and oxygen output is calculated as well as the relation between maximal power output and oxygen output. With regard to the indirect VO₂max measurement, the so-called field tests and laboratory ergometer tests may be used. However, accuracy of field test is not the subject of this paper and we will focus on the laboratory tests.

Over the years, several laboratory tests with different testing methodology were developed. The first submaximal cycle test was designed by Astrand and Ryhming in 1954, called the Astrand-Ryhming test (Astrand, 1954). This test extrapolates to the predicted maximal heart rate (HR) achieved from submaximal power output as its method of prediction. Another well-known submaximal cycle test is the Physical Working Capacity 150 (PWC₁₅₀) test, which uses power output achieved at a target HR of 150 beats per minute (bpm) as a prediction method (Campbell, 2001). Additionally, several treadmill tests were developed, with Balke (Balke, 1959) and Bruce (Bruce, 1972) being the most preferred. These tests use time to the point of exhaustion as a method of prediction. The aforementioned tests, however, were developed for diagnostics of cardio pulmonary disorders in sedentary population and were not designed for athletes. Therefore, choosing an inappropriate test could lead to inaccurate results, questioning its accuracy for specific population groups (Basset, 2000). Consequently, before choosing a suitable test, its validity should be adequately established. Various studies have attempted to verify indirect VO₂max prediction methods in laboratory conditions with the aim of assuring their validity in dissimilar population groups (sedentary population, adolescents and athletes) with constraining final results (Froelicher, 1974; Pollock, 1976, 1976 & 1982; Grant, 1999; Noonan, 2000; Chavda, 2013). Additionally, no study has been performed assessing largest athletic population, male soccer players, making findings on this topic rather scarce.

Therefore, this study was performed in order to compare the four abovementioned tests with CPET to determine whether they could be used for accurate VO₂max estimate in the case of male soccer players.

Materials and Methods

Participants

Sixteen healthy, midfield male soccer players took part in this study. The athletes competed in a national league, with ≥ 13 h of training per week over the last 8 y. They were recruited through a voluntary, open-access, online poll. Prior to the testing, all participants completed a questionnaire regarding their exercise and health history. Furthermore, they were made familiar with the ethical principles, potential health risks, aims and objectives and willingly signed a statement of consent prior to participating in the study. The entire research complies with the policy statement of the Declaration of Helsinki and was approved by the Institutional Ethics Committee. The subjects' body mass and height were assessed at the time of their first visit to the laboratory using a Seca 763 digital medical scale and stadiometer (SECA®, Hamburg, Germany). The athletes morphological characteristics were: age = 27.8 ± 1.8 y, height = 182.8 ± 2.7 cm and body mass = 81.5 ± 4.5 kg.

Exercise protocols

Experimental data collection was managed between 09:00 and 14:00 h during the preparatory autumn period. All tests were conducted in a controlled laboratory environment (steady 21 °C aired conditioned room, 43 % air humidity). Between each test, a period of 48 h was to elapse in order to ensure adequate physiological recovery. The Astrand and PWC₁₅₀ test were performed on a CX1 stationary bike (KETTLER® Ense-Parsit, Germany), while the Balke, Bruce and CPET test were carried out on a T170 motorized treadmill (COSMED® Rome, Italy). HR was collected continuously during the tests using a telemetric heart rate monitor (POLAR® Kempele, Finland) for all tests. The participants visited the laboratory on five separate occasions to perform test sessions in a predetermined order, at the same time of the day, to avoid circadian variance. During the experimental period, they were required to refrain from any systematic training activity 24 h prior to the each test and to consume their last meal, free of caffeine or nutritional supplements, 2 h before every test. All participants' had previous experience in ergometry testing.

Direct measurement

The preliminary CPET was assessed using a Quark PFT Ergo (COSMED® Rome, Italy) system by way of real breath-by-breath exchange data analysis during graded exercise testing (GXT) performed to the point of voluntary cancellation due exhaustion. The primary parameter measured was $\text{VO}_{2\text{max}}$ expressed in milliliters per kilogram per minute ($\text{ml}\cdot\text{kg}^{-1}\cdot\text{min}^{-1}$). Gas analyzers for oxygen (O_2) and carbon dioxide (CO_2) were calibrated for every test with a specific gas concentration (O_2 16.10 % and CO_2 5.20 % and NO_2 rest) according to the manufacturer's recommendation. Flow meter turbine calibration (3 L syringe) was performed on a test basis. The GXT protocol was comprised of three stages: inaction, exercise and recovery. The inaction stage lasted for 2 min to achieve a baseline HR. The exercise stage started at speed of $6 \text{ km}\cdot\text{h}^{-1}$ followed by speed rise of $1 \text{ km}\cdot\text{h}^{-1}$ every 2 min. The incline remained constant at 1%. When person had reached exhaustion and submits, the active recovery stage began lasting for 3 min at starting values. $\text{VO}_{2\text{max}}$ was determined by the following parameters: a respiratory quotient (RQ) ≥ 1.15 or a plateau of VO_2 in spite of a load increase (Howley, 1995). Quark PFT Suite monitoring software (COSMED®, Rome, Italy) was used to display data numerically and graphically in real time. Fluctuations of breath-by-breath data were minimized using six breaths smoothing and 30 sec averaging.

Indirect measurements

The indirect $\text{VO}_{2\text{max}}$ tests were performed in the following order: Astrand, Bruce, PWC₁₅₀ and Balke test. Test protocols and stage durations used were implemented to the nearest as earlier described (Astrand, 1954; Balke, 1959; Bruce, 1972; Campbell, 2001). The equations for $\text{VO}_{2\text{max}}$ calculation associated with the study tests, were designed for different population groups, scientifically validated and commonly cited. Accordingly, they were considered to be appropriate references for the observed population.

- Bruce test (Foster, 1984): $\text{VO}_{2\text{max}} = 14.8 - (1.379 \times T) + (0.451 \times T^2) - (0.012 \times T^3)$,
- Balke test (Pollock, 1976): $\text{VO}_{2\text{max}} = 1.444 \times T + 14.99$,
- Astrand-Ryhming test (Buono, 1989): $\text{VO}_{2\text{max}} = (0.00212 \times W + 0.299) / (0.769 \times \text{HR}_{\text{max}} - 48.5) \times 100$,
- PWC test (Rost, 1982): $\text{VO}_{2\text{max}} = 3.5 + 12 \times W_{\text{end}}$.

T stands for the total time of the test expressed in minutes and fractions of a minute, W is maximal workload achieved during test expressed in Watt's, HR_{max} stands for maximal heart rate achieved and W_{end} is maximal workload achieved during test divided per kilogram of body mass ($\text{Watt}\cdot\text{kg}^{-1}$). In order to be awarded with the T, W or W_{end} , subjects had to complete the commenced test stage in its entirety. All testing results achieved were manually logged in a data journal and evaluated in correspondence with the equations and methods discussed in the previous sections.

Statistical analysis

All data obtained were analyzed using MedCalc 12 (MEDCALC®, Belgium) statistical software and presented with a mean $\pm$ standard deviation (SD) and a 95 % confidence interval (CI). The Shapiro–Wilk test was used to observe normality of all measured data with normal distribution accepted. The one way ANOVA with

repeated measures was used to assess differences in mean $\text{VO}_{2\text{max}}$ between tests. Regression analysis with the coefficient of determination (R^2) was used to estimate the relationships between tests and the possibility of prediction of one result over another with CPET being dependent variable (Y) and other tests being the X variable. The Bland-Altman plot was used to compare measurement techniques and the possibility of replacement. The coefficient of variation (CV) was used to measure the relative variability of the data sets with CPET considered as the true value. The standard error of the estimate (SEE) was used to reflect the random error within the tests. Alpha levels for all hypothesis testing were set at $p \leq 0.05$ as the level of significance for statistical tests unless otherwise stated.

Results

An analysis of the data has found inconsistency of the $\text{VO}_{2\text{max}}$ values in the observed tests. The within-subjects characteristics with CPET values shown for comparison can be seen in Table 1. We observed highest differences related to measure in Balke test, followed by PWC₁₅₀ and Astrand-Ryhming test. When compared with CPET, highest similarity was present in Bruce test.

The pairwise comparison of one way ANOVA revealed highest $\text{VO}_{2\text{max}}$ mean difference between CPET and PWC₁₅₀ test (22.37, 95 % CI 17.49 to 27.25 and SEE = 1.49). The Astrand-Ryhming test (19.75, 95 % CI 15.04 to 24.45 and SEE = 1.43) and Balke test (11.75, 95 % CI 6.83 to 16.72 and SEE = 1.51) followed. The lowest difference and error of estimate was observed in Bruce test (6.26, 95 % CI 1.61 to 10.90 and SEE = 1.41).

Table 1. Physiological characteristics of participants ($n = 16$) $\text{VO}_{2\text{max}}$ ($\text{ml} \cdot \text{kg}^{-1} \cdot \text{min}^{-1}$) in all five tests

Protocols	PWC	Astrand	Balke	Bruce	CPET
Arithmetic mean *	31.25	33.87	41.84	47.36	53.62
SD	2.53	2.07	1.69	5.61	6.41
SEE	0.63	0.52	0.42	1.40	1.60
CV	8.10	6.11	4.04	11.85	11.95
95 % CI	29.90 to 32.60	32.77 to 34.97	40.94 to 42.74	44.37 to 50.35	50.20 to 57.0

* $p \leq 0.05$, n = number of participants, SD = standard deviation, SEE = standard error, CV = Coefficient of variation (SD*100/mean)

The regression equations, derived from the collected data can be observed in Table 2. The correlation was found to be significant ($p \leq 0.05$) only in the case of the Bruce test.

Table 2. Regression analysis results ($n = 16$)

Protocols	Astrand	Balke	Bruce	PWC
Regression Equation	$y = 3.84 + 1.47 x$	$y = -2.48 + 1.34 x$	$y = 23.07 + 0.65 x$	$y = 23.73 + 0.96 x$
R^2	0.23	0.13	0.32	0.14
Standard error for coefficients a and b	$a = 24.71$ $b = 0.73$	$a = 39.68$ $b = 0.95$	$a = 12.01$ $b = 0.25$	$a = 19.64$ $b = 0.63$
F ratio	4.07	2.00	6.55	2.33
p	= 0.06	= 0.18	= 0.02	= 0.15
SEE	5.84	6.21	5.48	6.15

n = number of participants, SEE = standard error of the estimate

The results of the Bland and Altman plot analysis can be seen in Figure 1. The highest divergence was between the PWC₁₅₀ test and CPET (Figure 1a), with 22.37 $\text{ml} \cdot \text{kg}^{-1} \cdot \text{min}^{-1}$ of VO_2 (CI 95 % 19.21 to 25.54, $p < 0.01$). The limit of agreement was 23.27 $\text{ml} \cdot \text{kg}^{-1} \cdot \text{min}^{-1}$ of VO_2 . The discrepancy between Astrand-Ryhming test and CPET (Figure 1b) was 19.75 $\text{ml} \cdot \text{kg}^{-1} \cdot \text{min}^{-1}$ of VO_2 (CI 95 % 16.70 to 22.80, $p < 0.01$). The range of limit of agreement was 22.45 $\text{ml} \cdot \text{kg}^{-1} \cdot \text{min}^{-1}$ of VO_2 . For Balke test (Figure 1c), mean difference was 11.76 $\text{ml} \cdot \text{kg}^{-1} \cdot \text{min}^{-1}$ of VO_2 (CI 95 % 8.56 to 14.99, $p < 0.01$). The limit of agreement was 23.59 $\text{ml} \cdot \text{kg}^{-1} \cdot \text{min}^{-1}$ of VO_2 . The Bruce test revealed lowest (Figure 1d) dissimilarity of the mean with 6.26 $\text{ml} \cdot \text{kg}^{-1} \cdot \text{min}^{-1}$ of VO_2 (CI 95 % 3.24 to 9.27, $p < 0.01$). The lower and upper limit of agreement was 22.17 $\text{ml} \cdot \text{kg}^{-1} \cdot \text{min}^{-1}$ of VO_2 .

Fig. 1. Plotting the $\text{VO}_{2\text{max}}$ mean differences between indirect methods and CPET with Bland and Altman method of approach. a) PWC₁₅₀ vs CPET; b) Astrand-Ryhming vs CPET; c) Balke vs CPET; d) Bruce vs CPET.

Discussion

This study is the first to examine the validity and existence of a connection between the observed four indirect $\text{VO}_{2\text{max}}$ prediction tests and direct $\text{VO}_{2\text{max}}$ assessment in the case of male soccer players. The strength of relationship was low, but nevertheless significantly correlated ($R^2 = 32\%$, $p < 0.05$) only in the case of Bruce test. Bruce test furthermore, demonstrated biggest similarity of the measurement dispersion ($CV = 11.85$), lowest standard error of the estimate ($SEE = 1.40$) and lowest mean difference ($6.26 \text{ ml} \cdot \text{kg}^{-1} \cdot \text{min}^{-1}$) compared with true CPET values. The highest differences were detected in the PWC₁₅₀ test followed by Astrand-Ryhming and Balke tests, making them a non-valid method of $\text{VO}_{2\text{max}}$ prediction in observed population. We consider a direct comparison of our study with any previous studies to be impeded due to the fact how no study was ever performed on male soccer players. However, a comparison with similar studies is possible.

In study by Grant et al. (1999), when observing Astrand-Ryhming test, $R^2 = 0.16$, $p < 0.05$ and a $CV = 9.19\%$ was reported when sedentary population (age 18–35 y) was tested, with tendency of the Astrand-Ryhming test to over-predict the real $\text{VO}_{2\text{max}}$. According to Kasch et al. (1984), the Astrand-Ryhming test is an inappropriate substitute for direct measurement of maximal oxygen intake in sedentary population (age range 30–66 y) due to high HR_{max} calculation inaccuracies (reported $R^2 = 0.34$ and $CV = 30$). The reasons for such a high error present can be linked to Astrand-Ryhming methodology. Original test recommends cycling at 50 revolutions per minute (rpm), hence McKay et al. (1976) suggests using higher values (80–100 rpm) for optimal $\text{VO}_{2\text{max}}$. Furthermore, extrapolation of the submaximal HR vs work rate regression line to predict the HR_{max} , which is then used to predict W_{max} , and finally to predict $\text{VO}_{2\text{max}}$ is highly age, body size and fitness level sensitive and open for errors in submaximal tests (Froelicher, 1974; Kasch, 1984). Hence, we can conclude how Astrand-Ryhming test is inappropriate method of $\text{VO}_{2\text{max}}$ estimation in case of athletes with tendency to under-predict actual oxygen consumption. We believe how this occurrence can be additionally explained by the higher fitness level, strong lower extremities and cycling adaptation of the athletes (Poole, 1996). Grant et al. (1999) tested a sedentary population (age 18 to 35 y) and observed a low connection between the Bruce test and CPET ($R^2 = 0.24$, $p < 0.10$ and $CV = 8.82\%$). Furthermore, this study used a different equation for indirect estimation compared to our research. Koutlianos et al. (2013) while testing a large group of athletes, reported $R^2 = 0.41$, $p < 0.05$ and a $SEE = 6.11\%$ for Bruce test, with overestimation of $\text{VO}_{2\text{max}}$. However, this study tested athletes from different sport disciplines and used the ACSM running equation to which the high validity coefficient and higher inter-subjects variation could be attributed. Froelicher et al. (1974) tested sedentary population and reported an

$R^2=0.76$ and a $SEE=4.71$ in Bruce test. Same study reported modest degree of relationship between actual and predicted VO_{2max} in Balke test ($R^2=0.64$ and $SEE = 3.95$). They explained their results by means of existence of high metabolic increments between the test stages. Pollock et al. (1976) reported $R^2 = 0.85$, $p \leq 0.05$ and $CV = 0.25\%$ for a sedentary elderly population in Balke test, using a self-developed nomogram. Although Pollock's nomogram was used in our study as well, we obtained no similar results in athletes. We consider subjects differences in age and fitness level as main inter-study bias (Poole, 1996). Our results were only consistent with Pollock to confirm under predictions of VO_{2max} . No study was found comparing the PWC_{150} test with CPET. Moreover, our results showed no statistical significance for the PWC_{150} test so we concluded that mentioned test is not a valid predictor of actual VO_{2max} in our observed group. The PWC_{150} test demonstrated highest under predictions of VO_{2max} values and error associated with it. Our observations can be contributed to the testing methodology (submaximal cycle test). An analysis of the data obtained by the Bland and Altman method of approach for limits of agreement between the CPET and predictive tests suggest ambiguous results in all four tests, with lowest mean difference and lowest CI 95 % observed in the Bruce test with statistical significance accepted. Accounting for our statistical goals, we consider the agreement interval sufficiently narrow for our purpose. Hamlin et al. (2012) revealed increases in limits of agreement caused by group homogeneity and higher participants' aerobic power, providing the basis for our obtained results.

For predictive test to be statistically significant, it is required to have a minimal $R^2 = 0.70$, $p \leq 0.05$ (Hamlin, 2012). Since R^2 depends on the variance and range of a group of scores, less variance will lead to low validity. In our study, the variance in all four tests was very low but only statistically significant in the case of Bruce test ($F - \text{ratio} = 6.55$, $p < 0.05$) where a real effect difference was observed. A study with a small sample size, such as ours, has a reduced chance of detecting a real effect and random error, thus being one of the possible reasons for our particular observations. However, our target population were midfield soccer players. Hence, if the target population is narrow, then the sample can be smaller but still be representative. Grant et al. (1999) explained that a low validity coefficient is influenced by high group homogeneity and that this could be the reason for a lower correlation rather than lack of connection. The same study reported that the validity of the test should, instead, be rather observed by the greater importance of the error present with an acceptance in the 10 % margin. Given that we tested athletes from one sport and an identical playing position, we observed a very high instance of homogeneity which could have affected the final results. Furthermore, lowest SEE and error of coefficients a and b was observed in the Bruce test. Ultimately, we observed that the previous studies tested heterogeneous groups (sedentary population with high age and fitness level range or athletes from different sport disciplines), directly affecting the validity coefficient. We believe, when a homogenous group of top athletes is tested, an underestimation of VO_{2max} values and a low validity coefficient should be expected when observed indirect prediction methods are used.

Furthermore, during direct VO_{2max} assessment, we employed GXT protocol, designed to reproduce real running conditions and objective VO_{2max} measurement (Jones, 1996). Previous studies, however, used the original Bruce, Balke or Astrand-Ryhming protocol during the direct assessment, affecting final results due to the displacement of actual VO_{2max} values. Additionally, the observed predictive tests use a different methodology for VO_{2max} estimation (cycle or treadmill ergometer) were higher VO_{2max} in treadmills then cycle ergometry tests is to be expected, additionally influencing the final outcome (Storer, 1990). Further reasons behind such inter-study differences could be linked to the equations associated with indirect VO_{2max} calculations (Bassett, 2000). The observed studies used different equations, making inter-study comparison hindered. We observed how only T, HR and load were taken into consideration when calculating the VO_{2max} . Foster et al. (1984) suggested age and gender as significant impact factors and reported a lack of forecast significance in treadmill tests when only exercise time as a predictor of VO_{2max} was used due to large metabolic increments. Furthermore, he reported increase in $R^2 > 0.20$ when gender, age and activity variables were added to standard equations. Koutlianos et al. (2013) demonstrated a connection between these "traditional" types of equations and inaccurate final results causing underestimation of VO_{2max} in athletes. As a result, we recommend modifications of the currently used equations for athletes.

Conclusion

Findings of this study enhance our understanding of indirect VO_{2max} prediction and may be applied to other similar studies worldwide. We do not consider these new findings trivial, as we do not support recommendations that one could actually substitute the direct VO_{2max} measurement. By observing the data obtained, we can conclude that statistically acceptable validity is present in the case of Bruce test and that the results obtained are sufficient for the studied population.

Acknowledgements

The authors would like to thank Ana Djordjevic for technical and language editing contribution. This study was supported by the Rudy Roth Doctoral Research Scholarship.

References

American Thoracic Society; American College of Chest Physicians. (2003). ATS/ACCP statement on

-
- cardiopulmonary exercise testing. *Am J Respir Crit Care Med*, 15:167(2): 211-277.
- Astrand, P., & Rhyming, I. (1954). A nomogram for calculation of aerobic capacity (physical fitness) from pulse rate during sub-maximal work. *J Appl Physiol*, 7(2): 218-221.
- Balke, B., & Ware, R. W. (1959). An experimental study of physical fitness of air force personnel. *U S Armed Forces Med J*, 10(6): 675-688.
- Bassett, D. R. Jr., & Howley, E. T. (2000). Limiting factors for maximum oxygen uptake and determinants of endurance performance. *Med Sci Sports Exerc*, 32: 70-84.
- Bruce, R. A. (1972). Multi-stage treadmill test of submaximal and maximal exercise. In: *Exercise Testing and Training of Apparently Healthy Individuals: A Handbook for Physicians*. American Heart Association, Dallas, 32-34.
- Buono, M. J., Roby, J. J., Micale, F. G., & Sallis, J. F. (1989). Predicting maximal oxygen uptake in children: modification of the Astrand-Rhyming test. *Pediatr Exerc Sci*, 1:278-283.
- Campbell, P. T., Katzmarzyk, P. T., Malina, R. M., Rao, D. C., Pérusse, L., & Bouchard, C. (2001). Prediction of physical activity and physical work capacity (PWC₁₅₀) in young adulthood from childhood and adolescence with consideration of parental measures. *Am J Human Biol*, 13: 190-196.
- Chavda, V. V., Mahavirsingh, H. R., Chintansinh, P., Pradnya, A. G., Hemant, B. M., & Chinmay, J. S. (2013). Predicted Maximal Oxygen Consumption (VO_{2max}) Values Obtained During the Maximal Treadmill Test Using Different Protocols, *NJIRM*, 4(2): 149-155.
- Foster, C., Jackson, S. A., Pollock, L. M., Taylor, M. M., Hare, J., Sennett, M. S., Rod, J. L., Sarwar, M., & Schmidt, D. H. (1984). Generalized equations for predicting functional capacity from treadmill performance. *Am Heart J*, 107 (6): 1229-1234.
- Froelicher, V. F. Jr., Brammell, H., Davis, G., Noguera, I., Stewart, A., & Lancaster, M. C. (1974). A comparison of the reproducibility and physiologic response to three maximal treadmill exercise protocols. *Chest*, 65(5): 512-517.
- Grant, A. J., Joseph, N. A., & Campagna, D. P. (1999). The Prediction of VO_{2max}: A Comparison of 7 Indirect Tests of Aerobic Power. *J Strength Cond Res*, 13(4): 346-352.
- Hamlin, M. J., Draper, N., Blackwell, G., Shearman, J. P., & Kimber, N. E. (2012). Determination of Maximal Oxygen Uptake Using the Bruce or a Novel Athlete-Led Protocol in a Mixed Population. *J Hum Kinet*, 31: 97-104.
- Howley, T. E., Bassett, R. D. Jr., & Welch, G. H. (1995). Criteria for maximal oxygen uptake: review and commentary. *Med Sci Sports Exerc*, 27(9): 1292-1301.
- Jones, M. A., & Doust, H. J. (1996). A 1% treadmill grade most accurately reflects the energetic cost of outdoor running. *J Sport Sci*, 14(4): 321-327.
- Kasch, F. W. (1984). The validity of the Astrand and Sjöstrand submaximal tests. *Phys Sports Med*, 12: 47-52.
- Koutlianos, N., Dimitros, E., Metaxas, T., Deligiannis, A. S., & Kouidi, E. (2013). Indirect estimation of VO_{2max} in athletes by ACSM's equation: valid or not? *Hippokratia*, 17(2): 136-140.
- McKay, G. A., & Banister, E. W. (1976). A comparison of maximum oxygen uptake determination by bicycle ergometry at various pedaling frequencies and by treadmill running at various speeds. *Eur J Appl Physiol Occup Physiol*, 35(3): 191-200.
- Noonan, V., & Dean, E. (2000). Submaximal Exercise Testing: Clinical Application and Interpretation, *Phys Ther*, 80: 782-807.
- Pollock, M. L., Dawson, G. A., Miller, H. S. Jr., Ward, A., Cooper, D., Headley, W., Linnerud, A. C., & Nomeir, M. M. (1976). Physiologic responses of men 49 to 65 years of age to endurance training. *J Am Geriatr Soc*, 24: 97-104.
- Pollock, M. L., Bohannon, R. L., Cooper, K. H., Ayres, J. J., Ward, A., White, S. R., & Linnerud, A. C. (1976). A comparative analysis of four protocols for maximal treadmill stress testing. *Am Heart J*, 92(1): 39-46.
- Pollock, M. L., Foster, C., Schmidt, D., Hellman, C., Linnerud, A. C., & Ward, A. (1982). Comparative analysis of physiologic responses to three different maximal graded exercise test protocols in healthy women, *Am Heart J*, 103 (3): 363-373.
- Poole, G. (1996). Canadian Society for Exercise Physiology. *Professional Fitness and Lifestyle Consultant Resource Manual*. Gloucester, Ontario.
- Rost, R., Hollmann, W. (1982). *Belastungsuntersuchungen in der Praxis. Grundlagen, Technik und Interpretation ergometrischer Untersuchungsverfahren* (English: *Stress tests in practice. Basics, Technics and Interpretation of testing methods*), Georg Thieme, Stuttgart.
- Storer, T. W., Davis, J. A., & Caiozzo, V. J. (1990). Accurate prediction of VO_{2max} in cycle ergometry. *Med Sci Sports Exerc*, 22(5): 704-712.