THE RELIABILITY AND SENSITIVITY OF INDICES RELATED TO CARDIOVASCULAR FITNESS EVALUATION

Jinzhou Yuan[1], Yibing Fu[2], Ruipeng Zhang[3], Xi Li[4] and Gongbing Shan[4]

[1]*Department of Exercise Science, Xi'an Physical Education University, Xi'an, China*
[2]*Shandong Provincial Hospital, Shandong University, China*
[3]*Shannxi Institute of Technology, Shannxi, China*
[4]*Department of Kinesiology, University of Lethbridge, Alberta, Canada*

Original scientific paper
UDC 796.015.68:611.1

Abstract:

Determining the recovery heart rate (RHR) index after various submaximal exercises is a popular and practical way for a population's cardiovascular fitness evaluation. These evaluations are based on the regression among RHR, load intensity, and maximal oxygen uptake. However, little work has been done to 1) explore the influences of body weight and height on these tests' reliability and sensitivity, and 2) compare the reliability among the tests. As a result, practitioners often choose tests without the appropriate criteria. This study researched the mentioned two aspects by evaluating 30 male college students via three common tests – *30 cm step test*, *40 cm step test* and *squat-up-down test.* The results showed that the reliability and sensitivity of the three tests were remarkably different. Adding body weight into the evaluation would improve both reliability and sensitivity. Considering all the influence factors, *30 cm step* test was the best one. These findings suggested that applying the relative RHR index (normalized by body weight) should be considered for a population's cardiovascular fitness evaluation in the future.

Key words: *reliability, sensitivity, cardiovascular fitness test, relative indices, comparison*

Introduction

Cardiovascular fitness is not only related to an individual's health, but also associated with one's fitness level. Therefore, both population health researchers and coaching practitioners have devoted great effort to exploring ways to measure and evaluate cardiovascular fitness. It is commonly known that the level of fitness is strongly dependent upon age, gender, race, anthropometry (body weight and body height) and training or exercise intensity. For decades, sport scientists have focused on finding reliable ways for its evaluation on various selected populations (Rowell, Taylor, & Wang, 1964; Bailey, Shephard, & Mirwald, 1976; Kline et al., 1987; Hartung et al., 1993; Maranhao & Farinatti, 2003). As a result, there are now numerous evaluation methods which exist and which show certain suitability for either the general public or athletes of different ages. All these methods can be divided into two types. One measures maximum aerobic capacity, i.e. maximum oxygen uptake (VO_{2max}). VO_{2max} is a direct measure of the maximum amount of oxygen used by an individual during exhaustive exercise (ACSM, 2005). VO_{2max} reaches a plateau when the work intensity increases to a certain level. A subject with good aerobic capacity has a higher VO_{2max} than an individual with poor aerobic capacity (Morrow, Jackson, Disch, & Mood, 2000). The direct measurement of VO_{2max} is currently the only reliable quantitative method for cardiovascular fitness evaluation. However, measurement requires costly equipment and exhaustive performance, which is hardly practical in population's fitness tests, where mass measurement is required. The alternative is indirect measurements, which are popular and widely used among practitioners. This kind of tests utilizes submaximal exercise instead of maximal exercise and has been used to estimate VO_{2max} since the 1960s (Jette, Campbell, Mongeon, & Routhier, 1976; Noonan & Dean, 2000; Quail, Vehrs, & Jackson, 1999; Rance et al., 2005). There are various forms of submaximal exercise used to estimate VO_{2max}: one-mile running/walking, step tests and the squat-up-down test (Zwiren et al., 1991). These estimations are based on the correlation and regression among the recovery heart rate (RHR) after exercise, load intensity, and VO_{2max}. The characteristics of all indirect methods are simple, economic, easy-to-apply and safe. Therefore, they are widely used by sport and fitness practitioners in their training evaluation (Hartung et al., 1993; McArdle et al., 1972). Although these methods are popular and practical, there are some problems related to their reliability and sensitivity. Previous studies have shown a wide variability in the reliability (Noonan & Dean, 2000). Some studies show that the reliability of indirect methods is exceptionally high. A study done by McArdle et al. (1972) using Queens College Step Test revealed that the reliability of test and re-test was .92 and validity was r=-.75 (correlation between RHR and VO_{2max}). While, Jette et al. (1976) reported that the reliability of Canadian home step test was only .79 with 102 subjects. The different results of reliability reported could be caused by using different subjects and/or a different workload of the test. These results indicate that the reliability of submaximal exercise test is strongly related to the type and sample group applied (Shephard, Cox, Corey, & Smyth, 1979). As such, an inappropriate selection of test-type may lead to either under-stressing or over-stressing an individual's fitness level so that the reliability of the test would decrease (Noonan & Dean, 2000). In practice, how to choose the proper test type for a physical evaluation should be the first concern for practitioners.

There are numerous submaximal exercise tests available depending on age, gender and fitness level. Step test is most popular, as it requires limited equipment (steps bench, heart rate monitor and rhythm sound equipment) and minimal space. The test has several sub-forms depending on the step bench height: Harvard Step Test (45 cm), Queens College Step Test (16.25" or 41.28 cm), Chester Step Test (12" or 30.5 cm), Canadian's home step test (12" or 30.5 cm), Chinese citizen step test (30 cm) and Japanese citizen step test (40 cm). Among these forms, most of them use the absolute RHR index – an index considering heart rate only, neglecting the influences of body weight and height – to evaluate cardiovascular fitness (GASC, 2003). By changing the step heights, the load time and rhythm, many kinds of evaluation protocols are established for different subjects (Shephard, 1980). One common characteristic of these protocols is to use a constant step height, but protocols vary from one to another. To the best of the authors' knowledge, there currently exists no comparison study to elaborate the influence of step height on the reliability and sensitivity of the tests for the same subject group. Furthermore, there is also no reference record showing the influence of different types of submaximal exercise tests on the reliability and sensitivity. Therefore, this study aims at closing the gap by comparing three widely applied tests: *30 cm step-test* (ST_{30cm}), *40 cm step-test* (ST_{40cm}) and the *squat-up-down test* (SUDT). One specific aim of this study was to determine which of the three tests would supply the highest reliability and sensitivity.

In addition to the type of test, body weight and body height also have significant influences on the validity of submaximal exercise tests (Ashley, Smith, & Reneau, 1997; Kline et al., 1987; Lloyd et al., 2003). Their influences depend strongly on the intensity of the test load. The method using a regression equation to estimate VO_{2max} considered the influence of body weight and body height (Buckley et al., 2004); but absolute RHR index did not (Hui & Cheung, 2004). Excluding body weight and height in absolute RHR index for fitness evaluation could lead to some uncertainty, which needs to be determined through studies. Therefore, the second goal of this study was to quantify the influences of a subject's body weight and height on the reliability and sensitivity of ST_{30cm}, ST_{40cm} and SUDT.

Method

Subjects

Thirty (N=30) male college students from Xi'an Physical Education University participated in the study. The test subjects had a mean age of 20.3 yrs±.2, average body weight of 64.0 kg±2.0 and average body height of 173.4 m±.9. They were randomly selected from the Department of Exercise Sciences based on the students' availability during the period of data collection. All the subjects were average college students without active sport training. The Human Subjects Research Committee of the Xi'an Physical Education University scrutinized and approved the test protocol as meeting the criteria of Ethical Conduct for Research Involving Humans, of the Natural Sciences & Engineering Research Council. All the subjects in the study were informed of the testing procedures, signed an approved consent form and voluntarily participated in the data collection.

Test procedures

It is well known that RHR is strongly influenced by temporary physical status (load level, activity level and pre- and post-movement). Therefore the following standardization of the test procedure was applied to assure the reliability of the raw data:

1) 30 times squat-up-down in 30 seconds

Before the test, each subject was asked to sit down quietly for 5 minutes, after which, the first heart rate (P_1) was collected using a 15-second interval, followed by 30 times of squat-up-down. During the movement of squat-up-down, the subject's arms were kept straight horizontally in an anteriorposterior direction. The squat-down had to reach the lowest body position and the squat-up was required to achieve the straight body position. Immediately after 30 times of squat-up-down, the second heart rate (P_2) was measured using the same interval length. The third collection of heart rate (P_3) was 45 seconds after the second measurement. The absolute RHR index (S) was determined using equations proposed by Cai and Wang (1993):

$$S=(P_1 + P_2 + P_3 - 200)/10.$$

2) Step tests

The same test procedure was applied for both ST_{30cm} and ST_{40cm} on different days. Five minutes of quiet rest was required before the step test. Directly after each rest, the subject was asked to do the step up and down movement on a bench (30 cm or 40 cm) at a rate of 30 steps/min for 3 minutes in total. A metronome was applied to control the step frequency. During the movement, the upper body was kept straight. One minute after the end of the step test, the first heart rate (f_1) was measured using a heart rate monitor for a 30-second interval. The second heart rate (f_2) collection was performed 2 minutes after the test. The third one (f_3) was measured 3 minutes after the test. Both absolute RHR indices (X_{30cm} and X_{40cm}) were calculated based on the following formula (GASC, 2003):

$$X = \frac{100 \times t}{f_1 + f_2 + f_3}$$

where *t* refers to the duration of the step test, which was 180 seconds. If a subject felt fatigue and could not keep up the pace of the test, he could stop the test and *t* would be recorded. In this study, however, all the subjects completed the 180-second test.

All test procedures were strictly controlled in terms of administration, organization and environmental conditions. Every subject performed only one test per day in the morning. Each test was repeated three times. Therefore, it required 9 days to finish the raw data collection for each subject.

Data processing and statistical analyses

The reliability coefficients of all indices were assessed using the following formula (Thomas & Nelson, 2001):

$$r = \frac{MS_S - MS_E}{MS_S} = 1 - \frac{MS_E}{MS_S} ,$$

in which MS_S is the mean of squares for subjects and MS_E is the mean of squares of error. All MS values were obtained through the two-factor repeated measures analysis of variance (ANOVA). The significance level of all tests was set at $p<.05$. The ANOVA analyses were performed using SPSS 10 for Windows (SPSS Inc, Chicago, IL, USA).

Results and Discussion

In order to elucidate both the reliability and sensitivity of the selected tests, both absolute and relative RHR indices were examined and discussed. Relative indices were obtained by normalizing the absolute RHR indices with body weight or body height. By comparing the results obtained, we were able to show how the intensity of loading and anthropometric data related to the evaluation of cardiovascular fitness. A summary of the three absolute RHR indices and their reliability coefficients are shown in Table 1.

The results showed that there were no significant differences ($p>.05$) among the three measurements, which indicated the reliability of the raw data. As for test reliability, it is found that ST_{30cm} is the most reliable among all the three tests, followed by SUDT, leaving ST_{40cm} to be the least reliable test. The reliability coefficients were .91, .84 and .82 for ST_{30cm}, ST_{40cm} and SUDT, respectively. The clearly higher reliability of ST_{30cm} could relate to the moderate load level of the test. During ST_{40cm} numerous subjects felt overloaded in comparison to the ST_{30cm} test. This subjective feeling was reflected by the difference of the reliability coefficient. Such a result

Table 1. Analysis results of the absolute RHR indices reliability

Type of test	Indices	Source	SS	df	MS	F-value	p-value	Reliability (r)
ST_{30cm}	X_{30cm}	Between subjects	8598.10	29	296.49	1.30	.28	.91
		Within subjects	1650.00	60	27.50			
		Total	10248.10	89				
ST_{40cm}	X_{40cm}	Between subjects	4216.77	29	145.41	1.42	.25	.82
		Within subjects	1559.33	60	25.99			
		Total	5776.10	89				
SUDT	S	Between subjects	456.47	29	15.74	.13	.88	.84
		Within subjects	152.11	60	2.54			
		Total	608.58	89				

Legend: SS - sum of squares, df - degrees of freedom, MS - means of squares is consistent with some of the previous studies. A recent study done by Buckley et al. (2004) showed that the load intensity was a dominant factor influencing the reliability of the cardiovascular fitness test. Although the load levels of ST_{30cm} and SUDT were about the same based on the subjective feeling of the test subjects, the reliability difference between the two tests was obvious. The reason could lie in the loading time and the methods of heart rate collection. The loading time was 180 seconds for ST_{30cm} and only 30 seconds for SUDT. As for heart rate collection, ST_{30cm} used post measurements only, while SUDT employed pre-test and post-test data collection. Since there is no literature available on this issue, further studies are needed to quantify the relationship between reliability and loading time, as well as the studies regarding heart rate values collection. Based on the results of Table 1, ST_{30cm} should be the best one of the three tests for the fitness evaluation of Chinese college students.

The other aspect of this study was to explore the sensitivity of the selected tests. To the best of the authors' knowledge, no study exists to explore the reliability of a test with consideration of its sensitivity. A common way to quantify the sensitivity of a test is to use the change of variance (CV). When different methods are applied for the same subject group, a higher CV obtained by a method means that the method will cause larger variations among the test subjects in comparison to other methods (Zhao, Huang, & Yuan, 1994). This can be interpreted as: the higher the value of CV, the more sensitive the test. A summary of the sensitivity of the three selected tests is shown in Table 2.

The results showed that SUDT had a highest sensitivity, followed by ST_{30cm} and ST_{40cm}. However, sensitivity can only contribute to the quality of a test, when the test reliability is high. As SUDT had the lowest reliability of the three tests, it is possibly not a suitable test to evaluate cardiovascular fitness when absolute RHR index is used.

As stated in the introduction, body weight and body height are significant influential factors of cardiovascular tests as they are closely related to the load intensity of a test. Neglecting these two factors in cardiovascular evaluation in previous studies (Cai & Wang, 1993; GASC, 2003) would overlook certain aspects. As such, the reliability and sensitivity of a test could be changed. For showing the influences of these two factors on the test reliability and sensitivity, this study examined the relative RHR index of the three selected tests. Two relative RHR indices were introduced: one was normalized by body weight (e.g. X_{30cm}/weight) and the other by body height (e.g. X_{30cm}/height). A summary of the relative indices is presented in Table 3.

The results showed that body weight had a significant influence on the test reliability while the influence of body height was at a minimal level. For the selected tests, ST_{30cm} using relative RHR index (X_{30cm}/weight) showed the highest reliability with a considerable increase from .91 to .95, raised by 4.52% in comparison to the absolute RHR index. Although ST_{40cm} had the largest increase rate in reliability after normalization by body weight – from .82 to .91, an increase of 10.35% - its reliability was still obviously lower than ST_{30cm}. As for the SUDT, this test possessed both the lowest reliability and increase rate after normalization by body weight.

Similarly, normalization by body weight had obviously increased the test sensitivity, while the influence of body height was negligible (Table 4). The results indicated that using the relative RHR index related to weight would improve the effectiveness of a test. In decreasing order, the sensitivity for the selected tests was SUDT, ST_{30cm} and ST_{40cm}.

Test reliability and test sensitivity are two dominant factors for researchers to consider when planning a test protocol. However, both factors are not equally weighted. Test reliability should be the primary consideration and test sensitivity can be the compensating aspect for improving the effec-

Table 2. Basic statistical results of the absolute RHR indices

Type of test	Indices	Measurements	RHR index		CV of RHR index	
			Mean	SD	%	Mean
ST_{30cm}	X_{30cm}	1	62.73	11.05	17.62	16.96
		2	64.90	10.94	16.86	
		3	63.67	10.44	16.40	
ST_{40cm}	X_{40cm}	1	51.83	7.76	14.97	15.39
		2	52.03	8.21	15.79	
		3	53.83	8.30	15.43	
SUDT	S	1	8.33	2.74	32.84	31.31
		2	8.37	2.45	29.24	
		3	8.17	2.60	31.86	

Legend: SD - standard deviation, CV - change of SD = (SD/Mean)×100%

Table 3. Analysis results of the relative RHR indices reliability (body weight is in kilograms and body height in meters)

Type of test	Indices	Source	SS	df	MS	F-value	p-value	Reliability (r)
ST_{30cm}	X_{30cm}/Weight	Between Groups	3.80	29	.13	1.18	.31	.95
		Within Groups	.41	60	.01			
		Total	4.21	89				
	X_{30cm}/Height	Between Groups	2864.10	29	98.76	1.32	.28	.91
		Within Groups	547.75	60	9.13			
		Total	3411.85	89				
ST_{40cm}	X_{40cm}/Weight	Between Groups	2.01	29	.07	1.61	.21	.91
		Within Groups	.39	60	.01			
		Total	2.40	89				
	X_{40cm}/Height	Between Groups	1391.85	29	47.99	1.46	.24	.82
		Within Groups	519.54	60	8.66			
		Total	1911.39	89				
SUDT	S/Weight	Between Groups	.15	29	.01	.17	.85	.87
		Within Groups	.04	60	.00			
		Total	.19	89				
	S/Height	Between Groups	150.51	29	5.19	.12	.89	.84
		Within Groups	50.31	60	.84			
		Total	200.82	89				

Legend: SS - sum of squares, df - degrees of freedom, MS- means of squares

Table 4. Basic statistical results of the relative RHR indices

Type of test	Indices	Measurements	RHR index		CV% of RHR index	
			Mean	SD	%	Mean
ST_{30cm}	X_{30cm}/weight	1	1.00	.22	21.87	21.53
		2	1.03	.21	20.50	
		3	1.02	.23	22.23	
	X_{30cm}/Height	1	36.19	6.34	17.53	16.96
		2	37.45	6.32	16.88	
		3	36.74	6.05	16.48	
ST_{40cm}	X_{40cm}/Weight	1	.83	.16	19.02	19.68
		2	.83	.17	20.16	
		3	.86	.17	19.87	
	X_{40cm}/Height	1	29.89	4.35	14.57	15.39
		2	30.03	4.77	15.87	
		3	31.07	4.83	15.56	
SUDT	S/Weight	1	.13	.05	34.26	34.62
		2	.14	.05	34.59	
		3	.13	.05	35.02	
	S/Height	1	4.81	1.57	32.69	31.31
		2	4.83	1.40	29.01	
		3	4.72	1.50	31.73	

Legend: SD - standard deviation, CV - change of SD = (SD/Mean)×100%

tiveness of a test. Based on this consideration, the best way for the cardiovascular fitness evaluation of male Chinese college students was ST_{30cm} using relative RHR index normalized by body weight. The current study initiated the normalization concept in cardiovascular fitness evaluation; future studies are needed to explore the validity of relative RHR index on various populations related to age, race and gender.

Conclusion

This study utilized reliability and sensitivity tests as well as normalization to quantitatively compare three common methods in cardiovascular fitness evaluation among college students. It was found that: 1) a subject's weight had a remarkable influence on the reliability and sensitivity of RHR index for cardiovascular fitness evaluation. The relative RHR index obtained through normalization by body weight would increase both reliability and sensitivity significantly; 2) *30 cm step* test was the best method among the three tests selected. These findings should supply valuable information for population health researchers and practitioners in their future fitness evaluation and/or future studies using different population samples.

References

ACSM - American College of Sports Medicine (2005). *ACSM's guidelines for exercise testing and prescription.* 7th ed. Philadelphia: Lippincott Williams & Wilkins.

Ashley, C.D., Smith, J.F., & Reneau, P.D. (1997). A modified step test based upon a function of subjects' stature. *Perceptual and Motor Skills, 85*, 987–993.

Bailey, D.A., Shephard, R.J., & Mirwald, R.L. (1976). Validation of a self-administered home test of cardiorespiratory fitness. *Canadian Journal of Applied Sport Science, 1*, 67–78.

Buckley, J.P., Sim, J., Eston, R.G., Hession, R., & Fox, R. (2004). Reliability and validity of measures taken during the Chester step test to predict aerobic power and to prescribe aerobic exercise. *British Journal of Sports Medicine, 38*(2), 197–205.

Cai, R., & Wang, L. (1993). Study on criterion of index to evaluate cardiovascular function. *Journal of Chinese Sport Science, 13*(6), 40–43.

GASC - General Administration Sport of China (2003). *The fitness test manual for Chinese Adults.* Beijing: People's Sport Publication.

Hartung, G.H., Krock, L.P., Crandall, C.G., Bisson, R.U., & Myhre, L.G. (1993). Prediction of maximal oxygen uptake from submaximal exercise testing in aerobically fit and nonfit men. *Aviation Space Environmental Medicine, 64*, 735–740.

Hui, S.S., & Cheung, P.P. (2004). Comparison of the effects of three stepping cadences on the criterion-related validity of a step test in Chinese children. *Measurement in Physical Education and Exercise Science, 8*(3), 167–179.

Jette, M., Campbell, J., Mongeon, J., & Routhier, R. (1976). The Canadian Home Fitness Test as a predictor for aerobic capacity. *Canadian Medical Association Journal, 114*, 680–682.

Kline, C.J., Porcari, R., Hintermeister, P., Freedson, A., Ward, R., McCarron, J., Ross, J., & Rippe, J.M. (1987). Estimation of VO_{2max} from a one-mile track walk, gender, age and body weight. *Medicine and Science in Sports and Exercise, 19*, 253–259.

Lloyd, L.K., Bishop, P.A., Walker, J.L., Sharp, K.R., & Richardson, M.T. (2003). The influence of body size and composition on FITNESSGRAM Test Performance and the adjustment of FITNESSGRAM Test Scores for skinfold thickness in youth. *Measurement in Physical Education and Exercise Science, 7*(4), 205–226.

Maranhao Neto, G., & Farinatti, P. (2003). Non-exercise models for prediction of aerobic fitness and applicability on epidemiological studies: Descriptive review and analysis of the studies. *Revista Brasileira de Medicina do Esporte, 9*(5), 304–314.

McArdle, W.D., Katch, F.I., Peschar, G.S., Jacobson, L., & Ruck, S. (1972). Reliability and interrelationships between maximal oxygen intake, physical work capacity and step-test scores in college women. *Medicine and Science in Sports and Exercise, 4*, 182–186.

Morrow, J.R., Jackson, A.W., Disch, J.G., & Mood, D.P. (2000). *Measurement and evaluation of human performance.* 2nd ed. Champaign, IL: Human Kinetics.

Noonan, V., & Dean, E. (2000). Submaximal exercise testing: Clinical application and interpretation. *Physical Therapy, 80*(8), 782–807.

Quail, J.W., Vehrs, P.R., & Jackson, A.S. (1999). Predicting VO_2max in African Americans and Whites with the 1-mile track jog test. *Measurement in Physical Education and Exercise Science, 3*(1), 1–14.

Rance, M., Boussuge, P.Y., Lazaar, N., Bedu, M., Van Praagh, E, Dabonneville, M., & Duché, P. (2005). Validity of a VO_2max prediction equation of the 2-km walk test in female seniors. *International Journal of Sports Medicine, 26*, 453–456.

Rowell, L.B., Taylor, H.L., & Wang, Y. (1964). Limitations to prediction of maximal oxygen uptake. *Journal of Sports Medicine, 19*, 919–927.

Shephard, R.J., Cox, M., Corey, P., & Smyth, R. (1979). Some factors affecting accuracy of Canadian Home Fitness Test scores. *Canadian Journal of Applied Sport Sciences, 4*(3), 205-213.

Shephard, R.J. (1980). The current status of the Canadian home fitness test. *British Journal of Sports Medicine, 14*(2), 114–125.

Thomas, J.R., & Nelson, J.K. (2001). *Research methods in physical activity.* 4th ed. Champaign, IL: Human Kinetics.

Zwiren, L.D., Freedson, P.S., Ward, A., Wilke, S., & Rippe, J.M. (1991). Estimation of VO_{2max}: A comparative analysis of five exercise tests. *Research Quarterly for Exercise and Sport, 62*(1), 73–78.

Zhao, Q.R., Huang, H., & Yuan, J. (1994). *Measurement and evaluation in physical education and sports.* Xi'an: Shann'xi People's Education Publication.

Submitted: May 5, 2008
Accepted: October 23, 2008

Correspondence to:
Assoc. Prof. Gongbing Shan, PhD
Director of Biomechanics Lab
Department of Kinesiology, University of Lethbridge
4401 University Drive, Lethbridge, Alberta, Canada, T1K 3M4
Phone: + (403) 329-2683,
Fax (403) 329-2709,
E-mail: g.shan@uleth.ca

Acknowledgements

This experiment was supported by the Research Center of the Xi'an Physical Education University. We would like to thank all the subjects who donated their time and expertise.

POUZDANOST I OSJETLJIVOST POKAZATELJA POVEZANIH S PROCJENOM KARDIOVASKULARNIH FUNKCIONALNIH SPOSOBNOSTI

Sažetak

Uvod

Opće je poznato da razina kondicijske pripremljenosti ovisi o dobi, spolu, rasi, antropometrijskim mjerama (tjelesnoj težini i visini) te intenzitetu treninga ili vježbanja. Već desetljećima sportski znanstvenici pokušavaju pronaći pouzdane načine za njezino vrednovanje na raznim selektiranim populacijama. Kao rezultat tih istraživanja, danas postoje brojne metode za utvrđivanje stanja kondicijske pripremljenosti koje su se pokazale primjerene ili za opću populaciju ili za sportaše različite dobi. Popularan i praktičan način za procjenu kardiovaskularnih sposobnosti jest utvrđivanje indeksa frekvencije srca u oporavku (RHR) nakon različitih sub-maksimalnih optrećenja. Ove procjene temelje se na regresiji između RHR, intenziteta opterećenja i maksimalnog primitka kisika (VO_{2max}). Ipak, vrlo malo znanstvenih istraživanja bilo je usmjereno na utvrđivanje: 1) utjecaja tjelesne težine i visine na pouzdanost i osjetljivost tih mjernih instrumenta i 2) na usporedbu koeficijenata pouzdanosti između testova. Kao posljedica toga, u praksi se izabiru i upotrebljavaju testovi bez odgovarajućih kriterija. Cilj je ove studije istražiti navedene aspekte utvrđivanjem stanja kondicijske pripremljenosti 30 studenata primjenom tri slična testa – step test na klupici visine 30 cm, step test na klupici visine 40 cm i čučnjevi.

Metode

Ispitanici su dobi 20,3±,2 godine, tjelesne težine 64,0±2,0 kg i visine 173,4±,9 centimetara. Svi ispitanici bili su prosječni studenti tjelesnog odgoja koji se nisu aktivno bavili sportom. Koeficijenti pouzdanosti za sve mjerne instrumente utvrđeni su formulom Thomasa i Nelsona (2001):

$$r = \frac{MS_S - MS_E}{MS_S} = 1 - \frac{MS_E}{MS_S},$$

u kojoj MSs predstavlja aritmetičku sredinu kvadrata rezultata ispitanika (varijanca pravih rezultata), a MS_E aritmetičku sredinu kvadrata pogreške (varijanca pogreške ili standardna pogreška mjerenja). Sve MS vrijednosti (sume kvadrata rezultata ili zajedničke varijance rezultata) dobivene su analizom varijance (ANOVA) za svaki paralelni test. Razina značajnosti za sve testove postavljena je na razinu od 0,05. Za obradu podataka korišten je program SPSS 10 za Windowse (SPSS Inc, Chicago, IL, USA).

Rezultati i rasprava

Rezultati istraživanja pokazali su da su pouzdanost i objektivnost ispitivana tri testa statistički značajno različiti. Normaliziranjem rezultata s koeficijentom tjelesne težine povećala bi se i pouzdanost i osjetljivost mjernih instrumenata. Uzimajući u obzir sve faktore koji utječu na pouzdanost i osjetljivost mjernih instrumenata, pokazalo se da je najbolji od testiranih testova *step test na klupici visine 30 cm*. Ovi rezultati ukazuju na činjenicu da bi se relativni indeks frekvencije srca u oporavku RHR (normaliziran tjelesnom težinom ispitanika) trebao u budućnosti koristiti za utvrđivanje kardiovaskularne sposobnosti ispitanika, osobito u populacijskim mjerenjima.