

Principal components analysis to evaluate ventilatory variability: comparison of athletes and sedentary men

R. R. T. Castro · M. Magini · S. Pedrosa ·
A. R. K. Sales · A. C. L. Nóbrega

Received: 3 August 2010 / Accepted: 7 October 2010 / Published online: 4 November 2010
© International Federation for Medical and Biological Engineering 2010

Abstract The present work quantifies, through principal components analysis (PCA) the relationships among the variability of breath-by-breath ventilatory parameters [minute-ventilation (VE), tidal volume (Vt), and respiratory rate (FR)] during a maximal progressive exercise test. The results show that the first and second eigenvalues of the covariant matrix contains almost 90% of the variables' variance possible to see through the PCA, which means that the problem can be reduced by a two-dimensional analysis. The results show a close similarity between the global variability in two groups test, athletes and sedentary (control). For the athletes group, the parameter Vt is responsible for the high VE variability values while in the sedentary group the FR is more relevant for VE variability. The result improves the knowledge about respiratory variability during exercise, showing that Vt's and FR's variabilities contribute in different ways to global ventilation variability during a maximal cardiopulmonary exercise test in athletes and sedentary men.

Keywords PCA · Variability · Ventilation · Comparison

1 Introduction

Homeostasis maintenance is a continuous process that depends on multiple feedback mechanisms which operate around set points. Thus, biophysical parameters variability is an integral aspect of physiological control and their analyses are useful so as to detect abnormal patterns which may indicate the worst prognosis in some diseases. For instance, low heart rate variability indicates bad prognosis in heart disease [10, 11]. Respiratory parameters could represent a relevant physiological phenomenon for this subject theme, although rarely addressed. The ventilation's oscillation in steady-state conditions is usually caused by instability of respiratory chemical control loops [1]. The concept that some diseases are diagnosed or prognosticated [5, 13, 18, 19] by the incidence of minute-ventilation (VE) oscillations in different situation yields the physiological description patterns and the conception of methods capable of detecting abnormalities in ventilatory fluctuations.

More specifically, the occurrence of VE oscillations during exercise, named periodic breathing, exhibits important prognostic information in patients with heart failure [12, 19]. Clinical practice shows that some patients may exhibit oscillations in VE exercise without complying with the strict criteria of periodic breathing [2, 11]

The absence of a universally accepted definition of periodic breathing and the need to manually identify this phenomenon are technical challenges to the widespread use of periodic breathing as prognostic criteria in heart failure [7]. Whether a graded quantitative assessment of periodic breathing would provide more detailed prognostic information than the presence or absence of ventilatory oscillations during exercise is a question that remains to be answered [7]. We have recently found that there is an inverse relationship between the time-domain VE variability during exercise and

R. R. T. Castro · S. Pedrosa
Exercise Physiology Laboratory, National Institute
of Traumatology and Orthopedics, Ministry of Health,
Rio de Janeiro, RJ 20230-024, Brazil

R. R. T. Castro · A. R. K. Sales · A. C. L. Nóbrega
Department of Physiology and Pharmacology, Universidade
Federal Fluminense, Niterói, RJ 24210-130, Brazil

M. Magini (✉)
Departament of Science and Technology, Universidade Federal
Fluminense, Rio das Ostras, RJ 28890-000, Brazil
e-mail: marcio.magini@pq.cnpq.br

the ejection fraction in heart failure patients without periodic breathing [2]. The ventilatory variability evaluation may add information in the assessment of breathing oscillations throughout the exercise. It is imperative to understand the physiologic pattern of any new prognostic variable so that pathological variations can be identified.

There are different available methods that permit biological signals analysis. Some of them can be useful to determine patterns in the general behavior of the measures and, from wavelets to statistical procedures, some of them can provide information on how the body works in terms of external or internal parameters [6, 14, 16, 17]. Indeed, in ventilation variability some methods are proposed but without a unique choice, usually due to the high analysis complexity which involves a strong correlation between different parameters collected in the experimental procedure [8, 15, 20]. This fact leads us to propose principal components analysis (PCA) to determine the correlation between different parameters obtained in a ventilation data collection.

Exercise training confers adaptations that are capable of altering the responses to a single exercise section. The adaptation of respiratory variability to physical training is still unknown. Considering the eminent importance of respiratory variability analysis and the unavailability of a method to perform such analysis, the present article aims at comparing the variability of ventilatory parameters [VE, respiratory rate (FR), and tidal volume (Vt)] on sedentary healthy men and athletes during a progressive maximal exercise test, using the PCA.

2 Methods

2.1 Volunteers

Eighteen male volunteers (9 sedentary and 9 athletes) were invited to participate in the study. All of them were considered healthy after medical evaluation which consisted of anamnesis and physical examination. None of them were smokers or had been in regular use of any kind of medication. Sedentary men were not involved in any regular physical activity. Athletes were professional soccer players from the same soccer team in Rio de Janeiro, Brazil.

2.2 Study protocol

All volunteers provided written informed consent to participate in the study after full explanation of the procedures and their potential risks. The investigation conformed to the principles outlined in the Declaration of Helsinki and have been approved by the Institutional Research Ethics Committee on Human Research.

The groups performed a maximal cardiopulmonary exercise test in a treadmill (Trackmaster 30 × 30, USA) following an individualized ramp protocol. Exercise test was interrupted whenever the perceived exertion reached grade 10 in a modified BORG scale. Cardiopulmonary exercise tests were performed with gas exchange and ventilatory variables being analyzed breath-by-breath using a calibrated computer-based exercise system (Ultima Cardio2 System, Medical Graphics Corporation, USA). The O₂ and CO₂ analyzers were calibrated before each test using a reference gas (12% O₂; 5% CO₂; nitrogen balance). The pneumotachograph used was also calibrated, with a 3 l syringe using different flow profiles. During each cardiopulmonary exercise test, a 12-lead electrocardiogram was continuously recorded (Cardioperfect, Welch Allin, USA) and heart rate registered. Oxygen consumption (VO₂); carbon dioxide production (CO₂), Vt, and respiratory rate or frequency (FR), were registered breath-by-breath. VE was calculated online (Breeze Software 6.4.1, Medical Graphics, USA). Specifically, we can understand the measures as follows: VE: volume of air exhaled from the lungs in 1 min, respiratory frequency (FR): the number of breaths in 1 min, and Vt: volume of air exhaled from the lungs in one breath. Thus, the sum of Vts from all breaths in a minute is equal to VE. In the case where Vt does not vary, VE is the product of respiratory frequency and Vt.

Principal components analysis

The principal components method used in this analysis consists in separate the ensemble in two distinct groups: athletes and control. This method was chosen based on its capacity to analyze the relationship between parameters that move together. This occurs in consequence of the instrumentation system which collects ventilator parameters. As main result, there is a redundancy of information which can be simplified in a single variable. Principal components method is a rigorous method for achieving this simplification. The method generates new variables called principal components. Each principal component is related to the original value through a linear combination of each

Table 1 Example of data used in PCA method

Athlete/control (data)	VE	Vt	FR
1	117.2	1.94	50.85
⋮	⋮	⋮	⋮
	47.7	2.3	32.09
⋮	129.4	1.88	50.85
	⋮	⋮	⋮
	47.4	2.9	36.36
N	128.4	2.94	41.1
	⋮	⋮	⋮
	35.7	1.54	40

experimental value. The analysis is focused on the relation of the first and second component, where the first principal component is a projection of each observation in a single axis in space with the maximum variance among all possible choices, whereas the second principal component is another axis in space, perpendicular to the first (Table 1).

The data contains the pieces of information as showed in Table 1. Where the columns are the ventilation parameters VE, Vt, and FR and the rows the values obtained from the equipment. Starting with this data for each group, the standard deviation was calculated for each parameter, this value was used in order to normalize the data. The covariance matrix and the linear combination of the variables were used to calculate the coefficients of the PCA

Table 2 Demographic and anthropometric data of volunteers ($n = 18$)

Variable	Control ($n = 9$)	Athlete ($n = 9$)	P value*
Age (years)	26 ± 6	22 ± 2	0.128
Weight (kg)	77.7 ± 11.0	70.6 ± 1.3	0.134
Height (m)	1.75 ± 0.06	1.75 ± 0.03	0.866
BMI (kg/m^2)	25.4 ± 3.04	23.05 ± 1.14	0.064
Peak VO_2 ($\text{ml}/\text{kg}/\text{min}$)	42.64 ± 5.62	47.81 ± 0.56	0.028

* Comparison between groups by student's t test

method, for this the data was standardizing subtracting the mean of each parameter from the same row of data and divide by each standard deviation. The covariance between two variables is given by:

$$\text{cov}(X, Y) = \frac{\sum_{i=1}^n (X_i - \bar{X})(Y_i - \bar{Y})}{n - 1}, \quad (1)$$

which result in a covariance matrix given by Eq. 2.

$$C = \begin{pmatrix} \text{cov}(X, X) & \text{cov}(X, Y) & \text{cov}(X, Z) \\ \text{cov}(Y, X) & \text{cov}(Y, Y) & \text{cov}(Y, Z) \\ \text{cov}(Z, X) & \text{cov}(Z, Y) & \text{cov}(Z, Z) \end{pmatrix} \quad (2)$$

The eigenvectors of the C are showed in the matrix of coefficients (M_{coef}), separated in two groups. The matrix of coefficients for athletes and control group are showed in Eqs. 3 and 4.

$$M_{\text{coef(athletes)}} = \begin{pmatrix} 0.6818 & -0.043888 & -0.73022 \\ 0.4554 & 0.80666 & 0.37672 \\ 0.5725 & -0.58939 & 0.56997 \end{pmatrix} \quad (3)$$

$$M_{\text{coef(control)}} = \begin{pmatrix} 0.6826 & -0.015905 & -0.73062 \\ 0.53864 & -0.66471 & 0.51771 \\ 0.49389 & 0.74693 & 0.44517 \end{pmatrix} \quad (4)$$

The scores were calculated from this matrix and the distribution of the first and second component along the

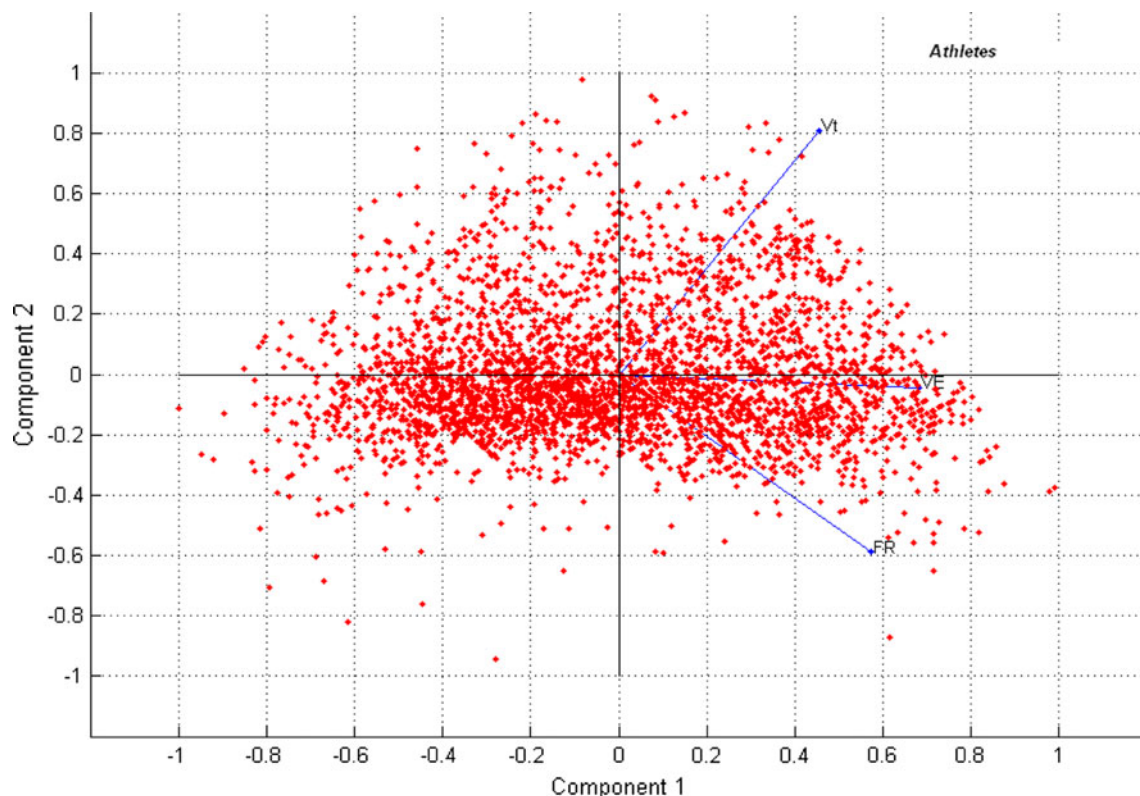


Fig. 1 Principal components analysis for the VE, Vt, and FR parameters in athletes group, where x -axis is the first component and y -axis is the second component for all values of time

axis generated the variability relation inner group. More specifically, the scores of each variable can be projected in a reduced coordinate system, in our case in two dimensional without loss of information, which contain the global behavior about the overall variance of the measures.

2.3 Statistical analysis

Statistical analysis was performed using the software Statistica 7.0 (Statsoft Inc, USA) for the demographic and anthropometric parameters. Descriptive data are presented as mean $\pm$ SD of the mean. Variables from the cardiopulmonary exercise tests showed normal distribution when analyzed by the Shapiro–Wilk's W test. Exercise variables in both groups were compared by paired bicaudal Student's t test. Significance was set at $P < 0.05$.

3 Results

All tests were considered maximal and reached a respiratory quotient greater than 1.10 (athletes: 1.34 ± 0.01 ; sedentary: 1.29 ± 0.13 ; $P = 0.38$). The demographic and

anthropometric characteristics of both groups are described in Table 2.

Figure 1 shows the VE, FR, and Vt values during the progressive maximal exercise test in athletes. It can be observed that FR and Vt contribute significantly more in time analysis. More specifically, the variances of first and second components for this group are 2.1365 and 0.85376, respectively. The third component assumes the value 0.0097691. It is important to notice that the first component has, approximately, 71.2% of the total data variance and the second component has 28.4%. Together, the first and second components are responsible for 99.6% of the variance.

These results show a lower data dispersion when this group (Fig. 1) is compared with the control group in Fig. 2. In sedentary men the Vt parameter has a small contribution for the group, which can be observed by the length of the vector Vt, the higher contribution for the variance came from the FR. The vectors are a projection of three independent components in a two-dimensional space. The coordinates of each vector in Fig. 1 are $VE = (0.6818, -0.0438)$; $Vt = (0.4554, 0.89066)$; and $FR = (0.5725, -0.5893)$ where the first value corresponds to the x -axis and the second value to the y -axis.

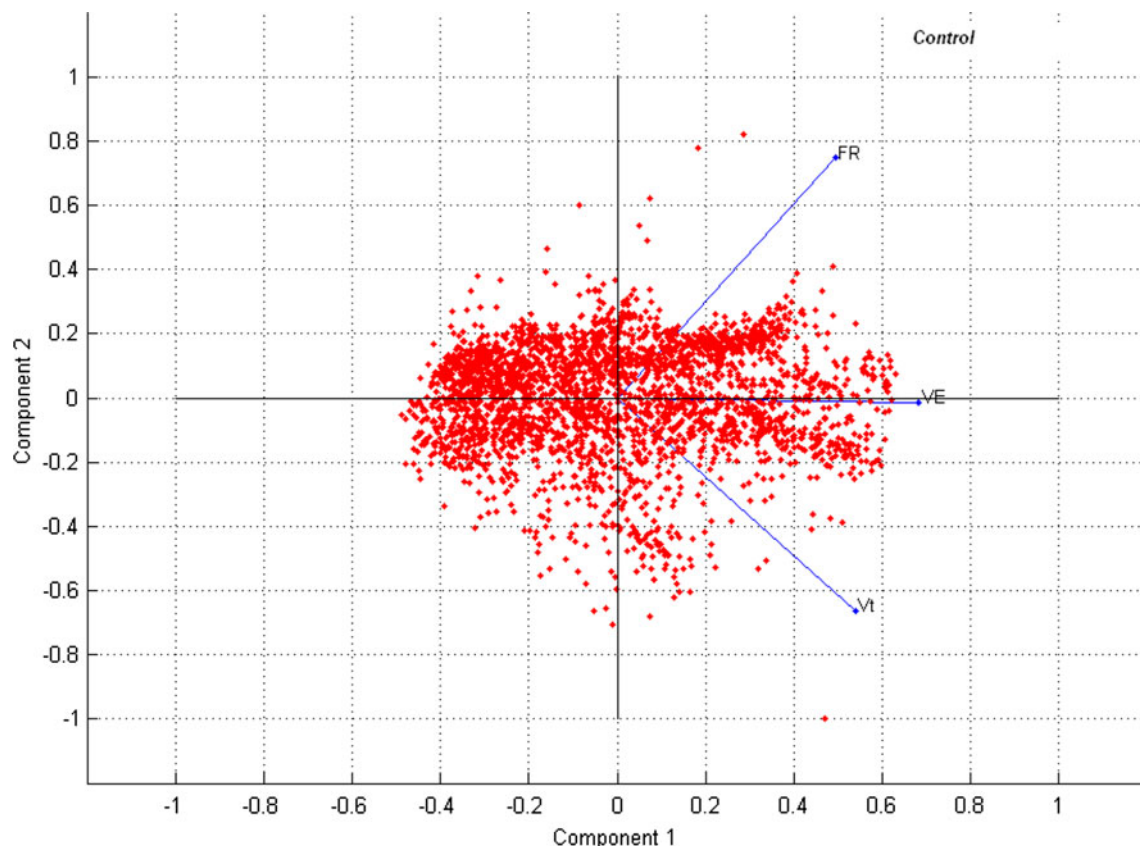


Fig. 2 Principal components analysis for the Ve, Vt, and FR parameters in control group, where x -axis is the first component and y -axis is the second component for all values of time

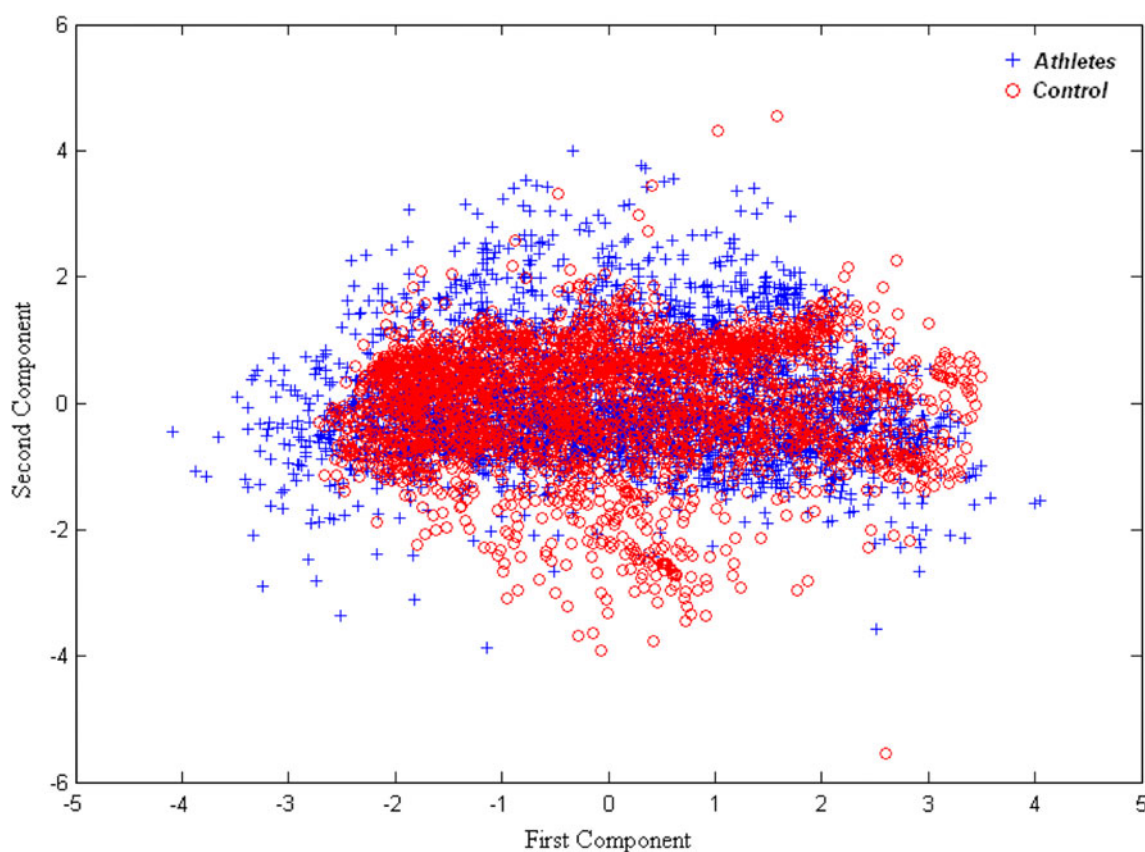


Fig. 3 PCA for ventilatory variability for both groups, the x -axis are the first component values and y -axis are the second component values

In the control group, the parameters change roles of importance. In this case, all parameters present almost the same contribution for the components value. It is necessary to register a more uniform dispersion around the zero. More specifically, the variance of the first and second component in PCA for the control group are 2.1137 and 0.85837. The third component assumes the value 0.027951. The first component is responsible for 70.5% of the total variance of the group followed by 28.6 and 0.9% from the second and third components, respectively. This fact promotes the result observed in Fig. 2, where data values distribution for the variance analysis through PCA shows a considerable number of points out of the boundary limits formed by the vectors VE, Vt, and FR. The vectors are a projection of three independent components in a two-dimensional space. The coordinates of each vector in Fig. 2 are $VE = (0.6826, -0.0159)$; $Vt = (0.5386, -0.6647)$; and $FR = (0.4938, 0.7469)$ where the first value corresponds to the x -axis and the second value to the y -axis.

Thus, the length of the vectors that represents the variables Ve, Vt, and FR defines their contribution for the data dispersion. Note that, the FR and Vt are perpendicular in Figs. 1 and 2, this fact is resulted from the physiological

meaning of Ve, which is the composition of these two measures. In Fig. 1, the vector Vt has the length bigger than the vector FR resulting in a higher contribution for the variation of Ve parameter. In Fig. 2 the roles are inverted, the length of the vector FR is bigger when compared with Vt length.

Figure 3 shows the PCA in two-dimensional space, this figure was generated from the union of the two first components figures, calculated separately, but designed in a single graph. In this case, it is possible to re-affirm that the VE, Vt, and FR dispersion is significantly higher in the control group, which is resulted from the values obtained in PCA. This result is expected but not easily observed. The PCA promotes a better level of understanding data dispersion. In this case, the control group presents higher global variability than the athletes group.

This cannot be statistically observed if the analysis is based on simple statistical procedure, as an example, in the box plots showed on Fig. 4. This figure show small differences between groups and the inversion of roles from the variables Vt and FR cannot be noted. The global analysis through PCA provides a general pattern's overview of the variables and its role for the group global dispersion.

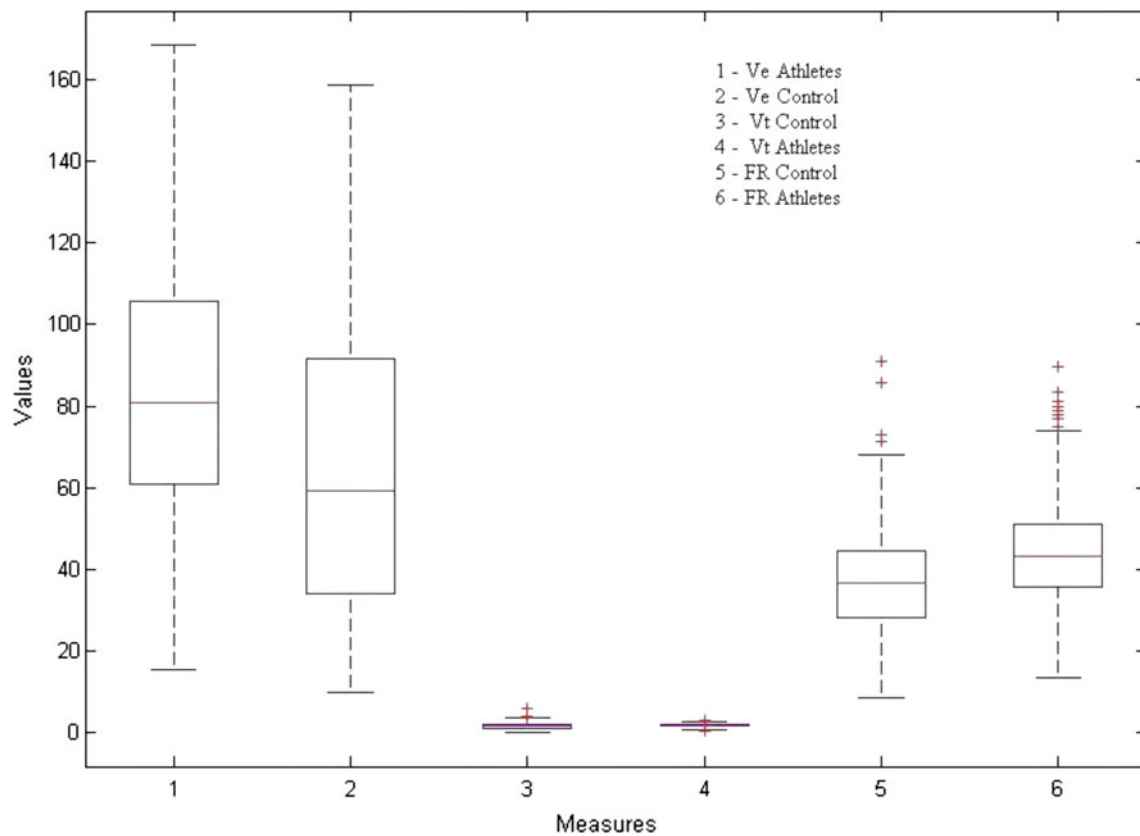


Fig. 4 Boxplot for the Ve, Vt, and RR values for two groups

4 Discussion

The response analysis of different biological parameters to exercise allows the precocious diagnosis of different pathological conditions. In this context, the excessive variability of ventilation parameters during exercise (periodic breathing) confers worst prognosis to patients with heart failure, regardless the presence of other classic prognostic parameters [10]. Periodic breathing seems to be the abnormal extreme of a more insidious process characterized by the inability of feedback loops to keep ventilation parameters varying around an accepted set point [3].

In fact, a myriad of abnormal ventilator variability responses during exercise seems to exist between normal pattern and periodic breathing [1, 2, 4]. During a progressive maximal exercise test, Vt, FR and thus, VE increase from rest to peak exercise. At low exercise intensity, the increase in VE is primarily accomplished by an increase in VE, while in more intense exercise the increase in FR is usually responsible for increases in VE [2, 9].

The variability of VE in athletes is secondary to a more pronounced variability in Vt, whereas in sedentary men the variability of FR seems to be the responsible for VE variability during exercise. These results could only be obtained due to the implementation of PCA, allowing a

better understanding on how Vt and respiratory vary and interact in order to interfere in VE variability in sedentary and trained men during a maximal progressive cardiopulmonary exercise test and showing the roles of the variables on exercise in two distinct groups.

References

1. Carley DW, Shannon DC (1988) A minimal mathematical model of human periodic breathing. *J Appl Physiol* 65(3):1400–1409
2. Castro RRT, Antunes-Correa LM, Ueno LM, Rondon MUPB, Negrão CE, Nóbrega ACL (2010) Ventilation variability inversely correlates to ejection fraction in heart failure. *Eur Respir J* 36(6)
3. Castro RRT, Antunes-Correa LM, Ueno LM, Rondon MUPB, Negrão CE, Nóbrega ACL (2010) Reversal of periodic breathing after aerobic training in heart failure. *Eur Respir J* 35(6): 1409–1411
4. Cherniack RM, Naimark A (1972) *Respiration in health and disease*. W.B. Saunders Company, Philadelphia
5. Corra U, Giordano A, Bosimini E, Mezzani A, Piepoli M, Coats AJ et al (2002) Oscillatory ventilation during exercise in patients with chronic heart failure: clinical correlates and prognostic implications. *Chest* 121(5):1572–1580
6. Ekkhakakis P, Lind E, Hall EE, Petruzzello SJ (2008) Do regression-based computer algorithms for determining the ventilatory threshold agree? *J Sports Sci* 26:967–976

7. Fleg JL (2010) Breathing not properly during exercise: prognostic implications in heart failure. *J Am Coll Cardiol* 55:1824–1825
8. Garrard CS, Das R (1987) Sources of error and variability in the determination of the anaerobic threshold in humans. *Respiration* 51:137–145
9. Hey EN, Lloyd BB, Cunningham DJ et al (1996) Effects of various respiratory stimuli on the depth and frequency of breathing in man. *Respir Physiol* 1:193–205
10. Hoyer D, Maestri R, La Rovere MT, Pinna DG (2008) Autonomic response to cardiac dysfunction in chronic heart failure: a risk predictor based on autonomic information flow. *Pacing Clin Electrophysiol* 31(2):214–220
11. La Rovere MT, Bigger JT Jr, Marcus FI, Mortara A, Schwartz PJ (1998) Baroreflex sensitivity and heart-rate variability in prediction of total cardiac mortality after myocardial infarction. ATRAMI investigators. *Lancet* 341:1470–1476
12. Leite JJ, Mansur AJ, de Freitas HF, Chizola PR, Bocchi EA, Terra-Filho M et al (2003) Periodic breathing during incremental exercise predicts mortality in patients with chronic heart failure evaluated for cardiac transplantation. *J Am Coll Cardiol* 41(12):2175–2181
13. Lorenzi-Filho G, Genta PR, Figueiredo AC, Inoue D (2005) Cheyne-Stokes respiration in patients with congestive heart failure: causes and consequences. *Clinics (Sao Paulo)* 60(4):333–344
14. Machado SM, Osório RAL, Silva NS, Magini M (2010) Bio-mechanical analysis of the muscular power of martial arts athletes. *Med Biol Eng Comp* 48:573–577
15. Myers J, Goldsmith RL, Keteyian SJ et al (2010) The ventilatory anaerobic threshold in heart failure: a multicenter evaluation of reliability. *J Card Fail* 16:76–83
16. Neto OP, Magini M (2007) Electromyographic and kinematic characteristics of Kung Fu Yau-Man palm strike. *J Electromyogr Kinesiol* 18:1047–1052
17. Paiva JPLM, Kelencz CA, Paiva HM, Galvão RH, Magini M (2008) Adaptive wavelet EMG compression based on local optimization of filter banks. *Physiol Meas* 29:843–856
18. Ribeiro JP, Knutzen A, Rocco MB, Hartley LH, Colucci WS (1987) Periodic breathing during exercise in severe heart failure: reversal with milrinone or cardiac transplantation. *Chest* 92:555–556
19. Sun XG, Hansen JE, Beshai JF, Wasserman K (2010) Oscillatory breathing and exercise gas exchange abnormalities prognosticate early mortality and morbidity in heart failure. *J Am Coll Cardiol* 55:1814–1823
20. Svedahl K, MacIntosh BR (2003) Anaerobic threshold: the concept and methods of measurement. *Can J Appl Physiol* 28: 299–323