

NEW SOFTWARE FOR THE IDENTIFICATION OF INDIVIDUALIZED RESPONSES TO PHYSICAL TESTS IN LARGE GROUPS OF SUBJECTS

Rafael Alkmin Reis¹, Denise Vaz de Macedo¹ and René Brenzikofer²

Laboratório de Bioquímica do Exercício, Universidade Estadual de
Campinas, Campinas-SP, Brazil¹

Laboratório de Instrumentação para Biomecânica, Universidade Estadual
de Campinas, Campinas-SP, Brazil²

The purpose of this study is to present a software developed to facilitate the visualization, identification and selection of subjects from large samples. The software is based on the use of scatter plots and allows the user, in an interactive way, to observe trends and patterns of the group's results, providing position of the individual within the group and showing the actual results of the subjects. The database used in this study consists of evaluations of strength, speed, endurance and body composition of 526 subjects, all beginners in the military career, collected three times along the year 2010. Because this software is easy to manage and allows the visualization of large amounts of data, it can be used by sports professionals to monitor athletes during competitive seasons and within the frame of talent identification processes.

KEYWORDS: large databases, athletes monitoring, talent identification.

INTRODUCTION: Physical evaluation is important and is applied as a routine procedure during training periods. Longitudinal monitoring of different physical fitness measurements can generate a large amount of data, especially when working with large samples. The identification of prominent subjects, the observation of relationships between variables, the selection of subgroups and the visualization of the effect of interventions are common initiatives in sports environment. These challenges receive special attention in the context of physical training and talent identification process (Gabbett et al., 2006).

The objective of this study is to present a software which allows easy viewing and managing of an extensive amount of data from large samples. This software may help in the observation of individual effects produced by physical training, assisting the talent identification processes in large groups of subjects.

METHODS: The database used in this software is composed of physical evaluations of 526 students, 18.5 ± 1.5 years old, from a preparatory school of the Brazilian Army. All subjects were evaluated three times, in February, May and October 2010. At each of these dates, nine physical tests were carried out. On the whole 21 variables were measured, corresponding to different physical abilities like strength, speed, endurance and body composition of the students. The software shows no limit either in the number of analyzed subjects or in the number of measured variables.

Initially all data were organized in Microsoft Excel and were checked for consistency between the recorded values and their relationship with the identities of the evaluated subjects. After this initial organization, all data were imported into Matlab® 7.0 which is the platform on which the software was developed. Despite having been developed in this environment, the software was compiled to an executable file and can now run on any computer using Microsoft Windows and does not require the installation of Matlab® 7.0. To be used first of all the software requests the user to choose two variables to be compared. With this pair of results, a traditional two-dimensional scatter-plot is performed. Each point in the chart is an evaluated subject. Examples are shown in Figures 1, 2 and 3. The originality of the software resides in the data management, the graph analysis and the easiness of manipulation. The Figures illustrate the output of the software on the screen. Besides the area reserved for the graph, the software provides on the right side buttons with specific functions that enable the user to interact with the graph.

1. *Group*: Allows the user to make the plot for a pre-defined subgroup (e.g. teams training for different modalities).
2. *Identify*: Enables, with a mouse click, the identification of the subject that corresponds to the point of interest.
3. *Search*: Draws a circle around the point corresponding to pre-selected subjects or group of subjects.
4. *Value*: Opens a window containing the real values of x and y coordinates of a point. This corresponds to results obtained by the subject in these evaluations.
5. *Identity Line*: Draws a line on the graph with slope equal to 1 and intercept equal to 0. This line represents the whole set of identical results in both tests if relevant.
6. *Trend Line*: Adjusts a linear regression line according to the criterion of least squares.
7. *Medians*: Draws two straight lines indicating the median values of each variable.
8. *Zoom*: Provides an enlarged view of a determined region of the graph.

RESULTS AND DISCUSSION: Three examples using the software will be presented. Intermittent sports, like volleyball, are known to include high intensity efforts as jumps and sprints interspersed by recovery periods. In Figure 1 we show a comparison between the 30-meter running velocity (m/s), X axis, and the distance reached in the horizontal counter movement jump (m), Y axis. Both variables were collected in February 2010.

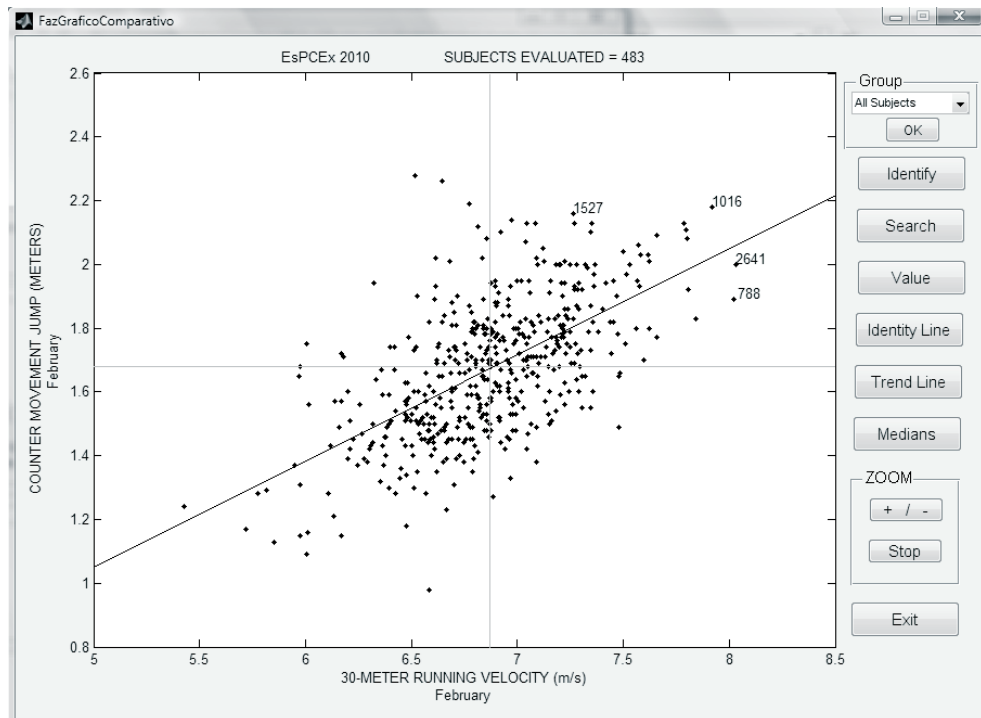


Figure 1: Layout of the software. Example of comparison between 30-meter running velocity (m/s), X axis, and horizontal counter movement jump (m), Y axis. The vertical and horizontal gray lines represent the medians for each variable. The sloping black line represents the linear regression line. The numbers in the graph represent the numerical identity of the subjects chosen with a mouse click. In this case, we can see the identities of some prominent subjects in both evaluations.

The advantage provided by the software resides in the interactive manipulation of the graph by the user. The 'Trend Line' function draws a linear regression line indicating the trend of the points. It makes evident that subjects with good performance in velocity tend to perform well in the jump test. This may be an important information in the context of volleyball.

With a click on the 'Medians' button, the software draws two gray lines at the median values for each of the variables. By doing this, we divide the graph area into four parts. This makes it easy to locate the region with the subjects more or less prominent in these two variables. With a single click on the button 'Identify', we discover the identity of the student related to the point of interest (Figure 1).

Another significant issue for the practice of volleyball is the subject's height and the upper limbs power. A possibility offered by the software is to localize desired subjects in the set of points. Figure 2 graphs the distance in medicine ball (3 kg) throw and subject height (m). For example, the identities of individuals 1527, 1016, 788 and 2641, the same subjects showing good performance in velocity and jumping already identified in Figure 1 can be searched.

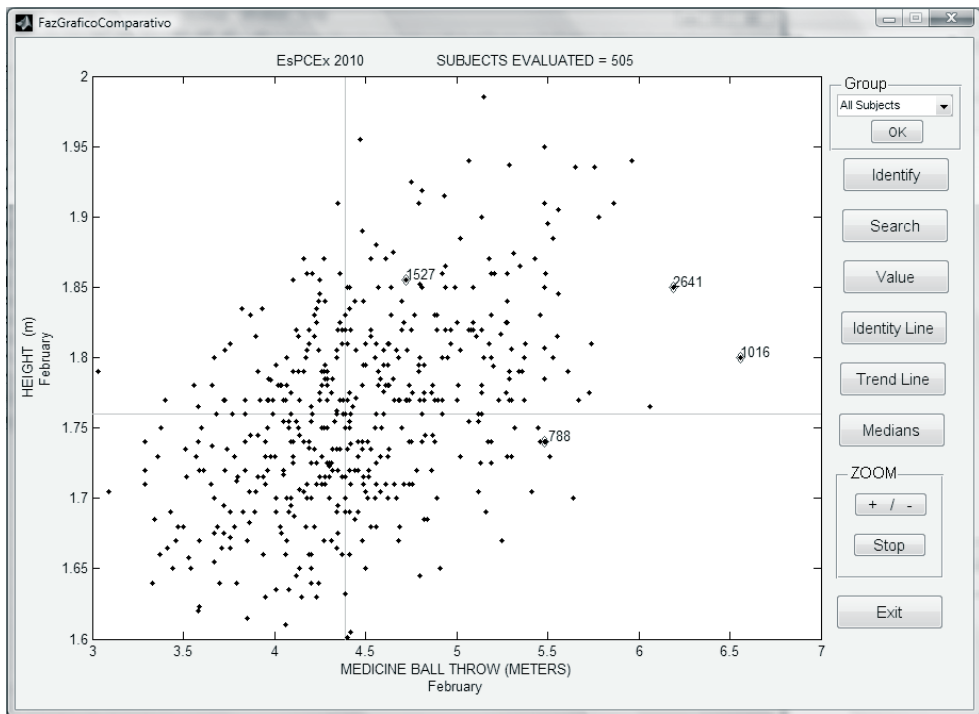


Figure 2: Comparison between the variables, distance of the medicine ball throw (m), X axis and subject's height (m), Y axis. The gray lines represent the medians for each variable. 'Search' is a function that looks for identities of interest and draws a black circle around them. In the figure above, students 1527, 2641, 1016 and 788 illustrate this example.

The analysis of Figure 2 shows that the four students identified on Figure 1 also deliver performance above the median in medicine ball throw. Subject 788 is the only student among those identified who happens to be below the median of the entire group under study. The example showed in Figure 3 illustrates the comparison between the time achieved on a 3000-meter run performed in February and the same variable measured in May. This type of comparison is of special interest because it allows visualizing the evolution of students during the training period between evaluations. By clicking on 'Identity Line', we draw a straight line with a slope equal to 1 and intercept equal to 0, which divides the graph diagonally. This line represents the set of identical results in both tests.

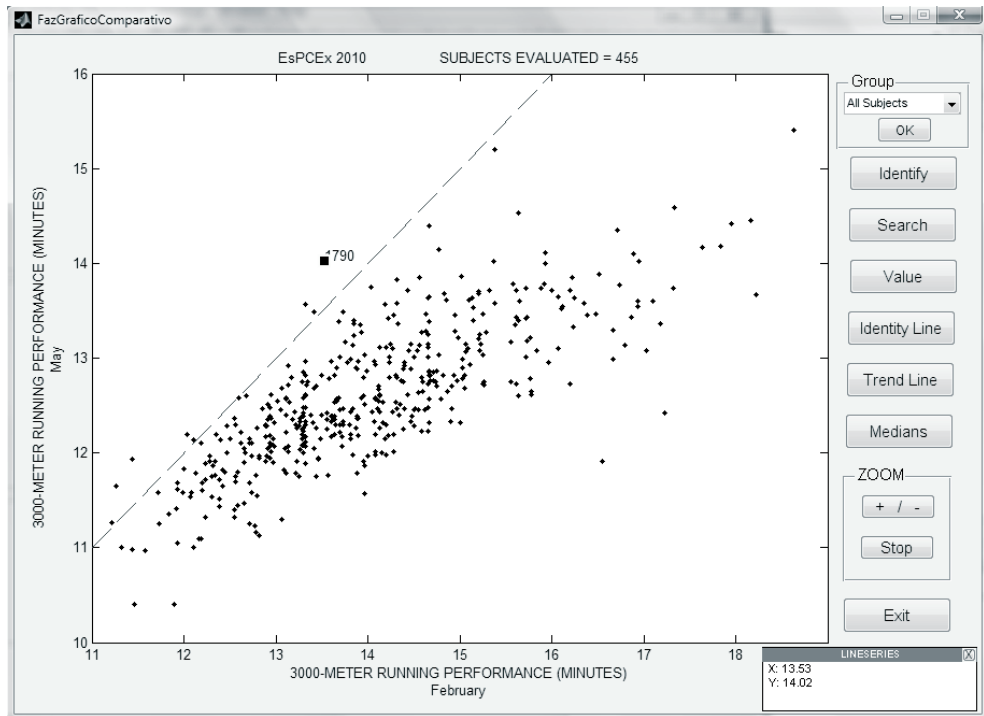


Figure 3: Comparison between variable time achieved on a 3000-meter race performed in February, X axis, and the same variable, but measured in May, Y axis. The dashed gray line represents the identity line. Value function opens a text box indicating the exact value obtained by the subject on both dates.

The points below the identity line represent the students who responded well to the physical training proposed, as they finished the 3000-meter run faster in May than in February. Note also how remarkable the performance improvement was for students who ran more slowly in February. But there are also a few students positioned above the identity line. These rare exceptions are individuals who, for some reason, performed the test more slowly in May than in February, for example student 1790 identified in the above Figure. In this case the subject performed the test in 13.53 minutes in February and 14.02 minutes in May, as shown in Figure 3 using the 'Value' function. The identification of students who did not respond effectively to the training process is essential for an individualized assistance to be provided.

CONCLUSION: The software presented in this study has been very useful for the identification of real effects of a specific intervention. It has provided assistance in talent identification process and selection of individual subjects from large samples. It is easy to manage and makes the information contained in large databases visible for a preliminary analysis. This facility makes it possible for any user, even without prior experience in computing or statistics, to observe trends and patterns about the group assessed in an interactive manner. Possible uses of this software in clinical analysis associated with some physiological biomarkers are also being studied.

REFERENCES:

T. Gabbett, B. Georgieff, S. Anderson, B. Cotton, D. Savovic & L. Nicholson. (2008). Changes in skill and physical fitness following training in talent-identified volleyball players. *Journal of Strength and Conditioning Research*, 20(1), 29-35.