The validity of incremental exercise testing in discriminating of physiological profiles in elite runners
The validity of incremental exercise testing in discriminating of physiological profiles in elite runners

A Legaz-Arrese¹, D Munguía-Izquierdo², LE Carranza-García¹, J Reverter-Masía³,
CG Torres-Dávila¹, RE Medina-Rodríguez⁴

¹Section of Physical Educations and Sports, University of Zaragoza, Zaragoza, Spain
²Section of Physical Educations and Sports, University Pablo de Olavide, Sevilla, Spain
³Section of Physical Educations and Sports, University of Lleida, Lleida, Spain
⁴Faculty of Sports Organization, Autonomous University of Nuevo Leon, Monterrey, Mexico

Received: 8 October, 2010
Accepted after revision: 25 January, 2011

The goal of this study was to determine whether traditional ergoespirometric incremental exercise testing carried out to the point of exhaustion could be useful in distinguishing the physiological profiles of elite runners that compete in races that lasted about 8 minutes versus those that lasted about 2 hours. Ten male marathon runners (performance time: 2:12:04, coefficient of variation (CV) = 2.33%) and 8 male 3000 m steeplechase runners (performance time: 8:37.83, CV = 2.12%) performed an incremental test on the treadmill (starting speed 10 km h⁻¹; increments, 2 km h⁻¹; increment duration, 3 min to exhaustion). Heart rate (HR), VO₂, and lactate concentrations were measured at the end of each exercise level. At maximal effort, there were no differences between the groups regarding VO₂max and maximal HR; however, the workload time, vVO₂max and peak treadmill velocity were significantly higher in the 3000 m steeplechase group (p<0.05). At submaximal effort, there were no significant differences between groups for VO₂, HR, or lactate. Our results show that this type of testing was not sufficient for discriminating the physiological profiles of elite runners who competed in middle-distance versus long-distance events (e.g. in the marathon and the 3000 m steeplechase).

Keywords: ergoespirometric, oxygen uptake, heart rate, lactate, endurance training incremental exercise testing, elite runners, workload time

Ergoespirometric incremental exercise testing carried out until the point of exhaustion is probably the most common laboratory test that is used to assess athletes competing in middle- and long-distance races. This test is a logical choice because it allows the assessment of many of the variables associated with performance in middle- and long-distance races [i.e. maximal oxygen intake (VO₂max), the minimum velocity corresponding to the VO₂max (vVO₂max), peak treadmill running velocity, aerobic and anaerobic thresholds, and running economy] in a relatively short amount of time (only 15–20 minutes). In addition, it also allows the definition of the different training zones [e.g. intensities that are lower than those that are required to reach the velocity of the aerobic threshold (vAeT), intensities between the vAeT and the velocity of the anaerobic threshold (vAT), intensities between the vAT and vVO₂max, and intensities higher than vVO₂max]. In addition, this test is necessary to determine other important variables, such as the maximal amount of time (Tlim) an athlete can spend at a specific intensity (e.g. Tlim at vVO₂max and the Tlim at a given percentage of VO₂max).

Corresponding author: Alejandro Legaz Arrese
Departamento de Fisiatría y Enfermería Universidad de Zaragoza
C/Domingo Miral S/N. 50009 Zaragoza, España
Phone: +34976761000; Fax: +34976761720
E-mail: alegaz@unizar.es
A large variety of protocols have been used to simultaneously determine variables associated with maximal and submaximal intensities during incremental exercise testing. In the protocols that require the measurement of lactate concentrations, a common method is for the athlete to perform 3 to 4 minutes of effort at each intensity level with an increase in the exercise intensity of 2 km.h⁻¹ and a resting period of 30 to 60 seconds at the completion of each intensity level to allow for blood sample collection (4, 14, 18, 24). Previous studies have shown that different variables are correlated with middle- and long-distance running performance with this type of testing, especially in heterogeneous groups (14, 18, 24, 30).

Usually, elite runners do not train to compete in both middle-distance (e.g. 1500 and 3000 m) and long-distance (e.g. 10000 m and marathon) events. In fact, elite runners are highly specialised in their training and work to improve their performance at a specific distance. For example, in 2009, there were no runners who were ranked in the top 50 athletes in the world in 3000 m or 3000 m steeplechase races and the marathon race (International Association of Athletics Federations, IAAF). Therefore, it is possible that different physiological profiles are conducive to success in middle- versus long-distance events. The aim of the present study was to determine if ergoespirometric incremental exercise testing until the point of exhaustion was useful in discriminating between the physiological profiles of elite 3000 m steeplechase and marathon runners.

Methods

Participants
A group of 10 male marathon runners [mean (SD) age 30.5 (4.1) years, body mass 60.7 (3.5) kg, height 1.72 (0.02) m] and 8 male 3000 m steeplechase runners [mean age 22.6 (3.4) years, body mass 66.1 (3.7) kg, height 1.77 (0.05) m] training to compete at top-level running events were selected based on their performances. Marathon runners generally performed continuous and interval training of long duration with a low intensity; in contrast, the 3000 m steeplechase runners routinely performed continuous and interval training sessions of shorter duration with a greater intensity.

Performance records in official races were obtained within three months before or after the time of experimental testing: for the group of male marathon runners, the time was 2:12:04 [Coefficient of variation (CV) = 2.33%], and for the group of 3000 m steeplechase runners, the time was 8:37.83 (CV = 2.12%). The performance of both groups was compared using Scoring Tables of the IAAF (26). Using a database of performances obtained at the world level, the IAAF assigns a score to each performance. This enables one to assess and compare different performances in different events. All athletes trained six or seven days a week (20–25 hours) during competition season.

The participants provided written informed consent to the experimental procedures after the possible benefits and risks of participation were explained to them. The study protocol was approved by the Research Ethics Committee of the Government of Aragón, Spain.

Designs
An experimental design was employed to determine if, a traditional ergoespirometric incremental exercise-test until exhaustion is useful in distinguishing the physiological profiles and training zones for marathon and 3000 m steeplechase.
Methodology
Each runner performed an incremental treadmill exercise test (Laufergotest LE-6, Jaeger, Germany). Subjects began running at an initial speed of 10 km·h⁻¹ at a slope of 1%, and their speed was progressively increased by 2 km·h⁻¹ every 3 min until the subject reached voluntary exhaustion. The exercise was interrupted for 30 s for blood sampling after completion of each 3-min workout. Blood samples obtained 10–15 s after the subject completed each exercise bout were measured by drawing capillary blood from the fingertip (Hitachi 4020, Boehringer, Ingelheim, Germany).

Gas exchange data measured during the last 30 s of each workout was used for the analyses. VO₂ values were measured using a Jaeger spiroergometer EOS-Sprint with a Hellige Servomed oscilloscope. VE was measured by means of a pneumotachograph that was specially designed to maintain linearity for high volumes; the exhaled CO₂ was measured using an infrared ray analyser, and the O₂, by means of a paramagnetic system (both using instruments from Jaeger). VO₂max was selected as the highest VO₂ value in the series of 30 s-by-30 s VO₂ values. VO₂max was considered to be reached according to primary and secondary traditional physiological criteria (10), namely the occurrence of a plateau in VO₂ despite an increase in running velocity, high levels of blood lactate concentrations (≥8 mmol·l⁻¹), elevated respiratory exchange ratio (r≥1.0), elevated heart rate (HR >90% of [220 – age]), and exhaustive perceived exertion (controlled visually and case-by-case). vVO₂max was considered to be the running velocity correspondent to the first stage that elicits VO₂max. In the majority of the runners in whom we did not observe a plateau of less than 2.1 ml·kg⁻¹·min⁻¹, the vVO₂max was calculated using the formula proposed by Kuipers et al. (12). Peak treadmill velocity was also estimated using the formula proposed by Kuipers et al. (12). The onset of blood lactate accumulation (OBLA) was detected by interpolation of the individually plotted treadmill velocity vs. blood lactate data.

Statistical analysis
Statistical analyses were performed with the Statistical Package for Social Sciences (SPSS), version 15.0. Data are expressed as mean (SD). The Student’s t-test and the Mann–Whitney U test were used to compare measurements between marathon and 3000 m steeplechase runners. Pearson and Spearman analyses were conducted to examine the relationships between the variables and running performance. Statistical significance was determined at the 0.05 level.

Results
The IAAF score was significantly higher in marathon (1157 ± 48) compared to 3000 m steeplechase runners (1105 ± 45) (p<0.05). No significant differences were found between groups regarding VO₂max (81.3 ± 4.0 vs. 80.5 ± 3.9 ml·kg⁻¹·min⁻¹, respectively, for marathon and 3000 m steeplechase runners) or maximal HR (190 ± 7 vs. 192 ± 6 lat·min⁻¹, respectively, for marathon and 3000 m steeplechase runners). Workload time was significantly longer in 3000 m steeplechase (1328 ± 68 s) than in marathon runners (1228 ± 41 s) (p<0.05). We observed that the 3000 m steeplechase runners had significantly higher vVO₂max values (22.3 ± 0.6 vs. 21.7 ± 0.5 km·h⁻¹, p<0.05) and peak treadmill velocities (22.7 ± 0.8 vs. 21.7 ± 0.5 km·h⁻¹, p<0.05) than marathon runners.
Fig. 1. Plot of oxygen cost expressed as ml.kg$^{-1}$.min$^{-1}$ in response to progressively increasing workload in athletes (solid circles = marathon runners; open circles = 3000 m steeplechase runners). The VO$_{2\text{max}}$ values that were measured at a velocity of 24 km.h$^{-1}$ only represent the parameters of four runners. This results in an artificial plateau in the VO$_{2\text{max}}$ values.

Fig. 2. Plot of oxygen cost expressed as ml.kg$^{-0.75}$.min$^{-1}$ in response to progressively increasing workload in athletes. (solid circles = marathon runners; open circles = 3000 m steeplechase runners). The VO$_{2\text{max}}$ values that were measured at a velocity of 24 km.h$^{-1}$ only represent the parameters of four runners. This results in an artificial plateau in the VO$_{2\text{max}}$ values.

Fig. 3. Plot of heart rate in response to progressively increasing workload in athletes (solid circles = marathon runners; open circles = 3000 m steeplechase runners). The HR values that were measured at a velocity of 24 km.h$^{-1}$ only represent the parameters of four runners. This results in an artificial plateau in the HR values.
There were no significant differences between groups in VO$_2$ (ml·kg$^{-1}$·min$^{-1}$) values obtained at submaximal intensities (Fig. 1). When VO$_2$ was expressed in ml·kg$^{-0.75}$·min$^{-1}$, the marathon runners demonstrated values that were significantly lower only at a speed of 14 km·h$^{-1}$ (p<0.05) (Fig. 2). No significant differences were found for submaximal values of HR (Fig. 3), lactate (Fig. 4), or OBLA intensity (19.21 ± 0.36 vs. 19.65 ± 1.2 km·h$^{-1}$, respectively, for 3000 m steeplechase and marathon runners).

As previously published, exercise lactate and HR were significantly associated with marathon running performance. This was true for exercise lactate in the treadmill velocity range of 10 to 18 km·h$^{-1}$ ($r^2$=0.73–0.91, p<0.05), and for exercise HR in the range of 10 to 20 km·h$^{-1}$ ($r^2$=0.65–0.76, p<0.05). Among 3000 m steeplechase runners, only workload time was significantly related to running performance ($r^2$=-0.71, p<0.05).

**Discussion**

The present study tested the differences in the physiological profiles of marathon and 3000 m steeplechase runners using incremental exercise testing until the point of exhaustion. The key finding was that the physiological results of both groups were similar. The main difference between groups was that the workload time, vVO$_{2\text{max}}$, and peak treadmill velocity were significantly higher in the 3000 m steeplechase runners than in the marathon runners.

The results from this study indicate that some of the parameters of the basic physiological profile were similar between athletes that were successful in races that lasted 8 minutes versus those that lasted 2 hours, despite the significant differences in exercise duration that are present in these events. The similarity between these parameters could explain the observation that during their sports careers, some athletes have been able to display an optimal performance in events that fall between the 3000 m distance and the marathon distance. However, if we exclude Haile Gebrselassie, who is one of the four all-time best runners in all distances ranging from the 3000 m to the marathon (excluding the 3000 m steeplechase) (IAAF, March 2010), an analysis using the ranking of Spain [Real Federación Española de Atletismo (RFEA), March 2010] and the world ranking (IAAF, March 2010) revealed that no runners have had comparable performances in distances that have such different durations. This analysis is
even more striking if we consider the simultaneous performance across different events. In 2009, none of the top 50 marathon runners, both within Spain and worldwide, were included (within their own ranking systems) among the top 50 in the 3000 m, 3000 m steeplechase and 5000 m events.

In our study, none of the runners competed simultaneously in the marathon and the 3000 m steeplechase or in any other intermediate distances that would allow a direct comparison of their performances in athletic competitions. Then, it was not possible to make a direct comparison between runners’ performances for either distances. However, when one considers the performance levels of both groups [even if one considers the slightly lower performance level (IAAF score) of the 3000 m steeplechase runners] along with the rankings discussed above, there is a compelling argument that, on average, it would be very unlikely for the group of the 3000 m steeplechase runners to match the performance of marathon runners in a marathon and vice versa. Based on this argument, the results of this study also suggest that incremental exercise testing cannot differentiate between the physiological profiles that are conducive to success in events that have significantly different durations (about 8 minutes and about 2 hours, in the cases of these two events). In addition, although each individual athlete could benefit from knowing the results of their exercise testing (for example, to identify individual training zones), different effort protocols are needed to discriminate between the parameters that are predictive of good performance for these two events of markedly different durations. Such effort protocols would need to be more robust at identifying and analysing certain variables and also include additional variables to provide enough power to discriminate between the physiological profiles of elite middle-distance and long-distance runners.

It is well known that the HR, VO$_2$, and serum lactate concentration values obtained in a traditional incremental exercise testing do not correspond to the values of those parameters at a steady-state, at least for exercise intensities above the anaerobic threshold (28). Some researchers have reported that for a particular submaximal intensity, the HR (25), the VO$_2$ (27) and lactate concentrate (5) increase with the duration of the effort. This increase was found to correlate with exercise intensity (i.e. the more intense the exercise, the larger the increase) (11, 25). A protocol involving gradual increases of 2 km.h$^{-1}$ is not likely to provide enough data density to discriminate between the physiological profiles of elite athletes. This fact is relevant if we consider the high sensitivity of the lactate concentration to small variations in exercise intensity near the anaerobic threshold (21, 31). Furthermore, it is known that the rest time between every increase in workload can influence many physiological parameters (9). It is not known whether alternative protocols would have been more successful at identifying important differences in the physiological profiles between the two groups of runners in our study. However, it is likely that the maximal effort parameters would not differ significantly if a different protocol were used, and if the parameters for the submaximal intensities were to change, they would likely change in the same direction for both groups of runners. Further studies are needed to confirm this hypothesis.

More importantly, these results support the concept proposed by Noakes (23) that an incremental exercise test is not the optimal test to evaluate the potential of athletes. During incremental exercise testing, the athlete does not know the expected duration of the effort, which is very different to the competition setting. In addition, during exercise testing, the intensity is gradually increased, so the athlete cannot regulate the intensity of the exercise. These constraints mean that this type of testing does not take into account or evaluate the importance of the brain in regulating exercise (22). Therefore, one can speculate that tests
that are more specific for a given distance, in which the duration and/or intensity of the exercise are kept constant, as well as tests that allow the athlete to regulate the intensity of the effort would be more appropriate for the assessment of elite runners.

With respect to the two events that were analysed in this study, a 3000 m test takes place at a velocity that is close to the vVO\textsubscript{2max} (13), whereas a marathon involves a velocity that is closer to the anaerobic threshold (18). Therefore, one could expect that during an incremental exercise test, the 3000 m steeplechase runners would exhibit a better physiological profile at the maximum intensity, whereas the marathon runners would exhibit a better physiological profile at submaximal intensities.

With respect to the parameters associated to a maximum intensity, only the peak treadmill velocity and vVO\textsubscript{2max} were higher in the 3000 m steeplechase runners than in the marathon runners. Because the two groups had similar VO\textsubscript{2max} values, these results are probably a consequence of the recruitment of more motor units and/or an increase in the contractile capacity of the muscles of 3000 m steeplechase runners, which allow them to better adapt to the conditions of anaerobic metabolism and/or other factors associated with fatigue (22). The similarity between groups with regard to VO\textsubscript{2max} is not surprising if we consider that at both distances, the higher VO\textsubscript{2max} value associated with maintenance of the %VO\textsubscript{2max} in a competition event results in a greater running velocity. An ergospirometric incremental exercise testing completed until a point of exhaustion is reached to determine VO\textsubscript{2max}. This study, like others (20), demonstrates that a high VO\textsubscript{2max} is required for optimal performance in middle- and long-distance events. However, as previously reported, this variable does not differentiate between elite runners competing in races longer than 1500 m (20), and unlike other variables, such as skinfold thicknesses (15) and left ventricular diameters at end-diastole (19), longitudinal studies have demonstrated that the changes in running performance are not related to changes in VO\textsubscript{2max} (17). In light of the previously presented arguments, one could speculate that the determination of the Tlim at vVO\textsubscript{2max} or at a specific absolute velocity close to the vVO\textsubscript{2max}, the determination of the %VO\textsubscript{2max}, and/or the velocity attained for a duration of about 8 minutes would allow investigators to identify differences in the physiological profiles of 3000 m steeplechase and marathon runners. It should also be considered that performance on such tests is associated with a greater interindividual variability (8). Thus, one can assume that these tests have a larger capacity to discriminate between the physiological profiles of runners who compete in events of different distances and also to discriminate between homogeneous groups of athletes who compete in medium-distance events. Interestingly, Basset et al. (2) observed that the VO\textsubscript{2max}, the vVO\textsubscript{2max} and VO\textsubscript{2} at submaximal intensities were equivalent between medium- and long-distance runners, but the medium-distance runners developed a 23% higher Tlim at vVO\textsubscript{2max}, probably as a result of their greater anaerobic capacity, among other reasons. It also would be logical to consider incorporating specific tests to assess anaerobic metabolisms and the manifestation of force in 3000 m steeplechase runners. In this context, one should consider that although the mean velocity in a 3000 m steeplechase race is close to that of the vVO\textsubscript{2max}, analysis of the optimal strategy for events of similar duration shows that the velocity should be higher during the end of the race (29). Therefore, during the most decisive part of the race, a 3000 m steeplechase runner moves at a velocity that is higher than the vVO\textsubscript{2max}, which requires greater forces and also involves anaerobic metabolism. In fact, we have previously observed elevated lactate concentrations in elite 3000 m runners after official competitions (17.2 mmol.l\textsuperscript{-1} in men and 16.3 mmol.l\textsuperscript{-1} in women, respectively) (16).
With respect to parameters measured at submaximal intensities, the results showed that the values of $\text{VO}_2$/intensity, HR/intensity, and lactate/intensity were similar between 3000 m steeplechase and marathon runners. Considering the arguments previously presented, one could hypothesise that the determination of these parameters during a longer effort of duration, such as during those obtained at the traditional velocity of maximal lactate steady-stable protocol (VMLSS) (3), would reveal a physiological profile that would be favourable to the marathon runners. The determination of the Tlim at the VMLSS or at a specific absolute velocity close to the VMLSS, the determination of the $\%\text{VO}_{2\text{max}}$, $\%\text{VMLSS}$, and/or velocity developed for a two-hours duration would be probably more specific tests to evaluate marathon runners. It should be kept in mind that the performance of these types of tests is subject to a large degree of inter-subject variability (6, 7, 25) and is also sensitive to the effects of training (5). Interestingly, it has been shown that during a two-hour effort, swimmers select a strategy that allows them to develop the highest possible velocity and keep their serum lactate concentration stable (1). In another interesting study, significant individual differences associated with marathon running performance were observed between running economy in the pre-fatigue condition and after running 30 km (6). The results of both of these studies suggest that these kinds of tests can be more sensitive in discriminating between the physiological profiles of athletes that compete in events of different duration and between homogeneous groups of athletes who compete in long-distance events.

**Conclusions**

The results of this study show that traditional incremental exercise testing is not sufficient in differentiating the physiological profiles of elite runners that compete in middle-distance versus distance events that range in duration from 8 minutes to 2 hours. Specifically, our results demonstrate that although this test is valid for the determination of $\text{VO}_{2\text{max}}$, this variable does not differentiate between top-class middle- and long-distance runners. Additionally, our results suggest that this test might not be best suited for the determination of submaximal variables; therefore, the training prescribed based on the results may be inaccurate. This finding suggests that more specific tests are required to determine the physiological profile that is conducive to success for each particular event distance. Future studies should be performed to further expand these results to other types of continuous efforts (e.g. swimming, rowing, cycling) and to other effort protocols. In addition, further studies should be conducted to identify tests that can better assess the potential and training status of athletes that compete in track events of different distances.

**REFERENCES**

