RESEARCH ARTICLE

European Journal of Sport Sciences
www.ej-sport.org

Exercise Intensity in Track and Level-grade Treadmill Running: A Cross-over Longitudinal Study in Well-trained Athletes

Aldo F. Longo, Enrique O. Prada, Marcelo L. Cardey, and Gustavo D. Aquilino

ABSTRACT

Heart rate is a conventional indicator of exercise intensity. Diverse studies have reported results of the comparison between the heart rate responses attained during running overground and on a treadmill; non-unanimous conclusions have emerged. The intention of this study was to compare the exercise intensity through heart rate between progressive running tests performed on track and level-grade treadmill. The heart rate responses of twelve highly trained male athletes were analyzed (Age = 24.3±2.7 years). The running protocol had initial and final speeds of 11 km·h^{-1} and 18 km·h^{-1}, and increments of 0.5 km·h^{-1} every 200 m. Two tests were performed: on an outdoor 400 m track, and a level-grade motorized treadmill under laboratory conditions. An innovative data analysis approach was proposed, by using a linear mixed-effects model, with the Test and Speed stage and their interaction as fixed factors, and the Subject as a random factor; a suitable correlation structure was also specified. The statistical significance level was set at $p < 0.05$. The difference between tests was not significant ($F = 0.06$, $p = 0.81$). The interaction effect between the Test and Speed stage was also not significant ($F = 1.32$, $p = 0.19$). Exercise intensity as measured by heart rate showed similar mean responses in track and level-grade treadmill running across a wide range of speeds in well-trained athletes.

Keywords: Exercise intensity; Heart rate; Progressive run; Track and level-grade motorized treadmill.

Published Online: November 02, 2022

ISSN: 2796-0048

DOI: 10.24018/ejsport.2022.1.6.37

A. F. Longo*
Exercise Physiology Laboratory, National Center of High Performance Athletics (CeNARD), Buenos Aires, Argentina.
(e-mail: alongo@deportes.gov.ar)

E. O. Prada
Exercise Physiology Laboratory, National Center of High Performance Athletics (CeNARD), Buenos Aires, Argentina.
(e-mail: eprada@deportes.gov.ar)

M. L. Cardey
Exercise Physiology Laboratory, National Center of High Performance Athletics (CeNARD), Buenos Aires, Argentina.
(e-mail: chelocardey@hotmail.com)

G. D. Aquilino
Exercise Physiology Laboratory, National Center of High Performance Athletics (CeNARD), Buenos Aires, Argentina.
(e-mail: gustavoaquilino@hotmail.com)

Corresponding Author

I. INTRODUCTION

Field and treadmill running tests are commonly performed for submaximal and maximal testing. Running protocols have been designed to be carried out outdoors and under laboratory conditions, generally involving the use of a track and a motorized treadmill, respectively. The results of these evaluations are used for diagnostic testing, to adjust exercise prescriptions, and to monitor training adaptations (American College of Sports Medicine, 2014; Gibson *et al.*, 2019).

Treadmill running is frequently assumed to replicate overground running (Miller *et al.*, 2019). This running condition facilitates the accurate adjustment of the kinematic parameters and offers a suitable environment for workload monitoring. Motorized treadmills have been extensively used for the clinic, training, athletic performance testing, and research purposes; a variety of outcome measures has been studied: physiological measures (heart rate, oxygen uptake, and blood lactate concentration), perceptual measures (rating of perceived exertion and preferred running speed), and performance measures (time trial, maximal sprint speed, maximal aerobic speed) (Miller *et al.*, 2019).

Heart rate is a conventional indicator of physical workload, and has been used in sports medicine and related fields as a key physiological measure to estimate exercise intensity (American College of Sports Medicine, 2014; Kenney *et al.*, 2019). Heart rate probably represents the most popular method to prescribe aerobic exercise intensity (Reuter & Dawes, 2016). The measurement of the cardiac response by heart rate monitors is typically used to quantify the extent of physical workload in training and testing procedures, given its non-invasive nature and easy implementation. The examination of the heart rate behavior across different levels of graded exercise allows for a more accurate selection of workload intensities according to specific aims. Furthermore, mathematical equations have been generated to calculate the target heart rate as a function of the desired percentage of exercise intensity, and physical activity intensity levels have also been defined based on the heart rate reserve and the maximal heart rate (American College of Sports Medicine, 2014; Kenney *et al.*, 2019).

Diverse studies have reported results of the comparison between the HR responses attained during running overground and on level-grade treadmill (Brookes *et al.*, 1971; Ceci & Hassmén, 1991; Chu *et al.*, 2010; Di Michele *et al.*, 2009; Jones & Doust, 1996; Köklü *et al.*, 2020; Kunduracioglu *et al.*, 2007; Maksud *et al.*, 1971; McMurray *et al.*, 1988; Miller *et al.*, 2019; White *et al.*, 1998; Yngve *et al.*, 2003). Athletes and non-athletes have been evaluated, and constant and progressive incremental protocols have been proposed. Statistical methods ranging from the simple paired t-test to multi-way analyses of variance accounting for repeated measures have been applied (Montgomery, 2013; SAS Institute Inc., 1999; Van Belle *et al.*, 2004). Unfortunately, non-unanimous conclusions have emerged.

A crucial issue in research studies aiming to contrast the exercise intensity between overground running and treadmill running is to avoid or minimize the risks of intensity bias, environmental bias, and statistical bias. This is, to ensure that the physical workload is effectively matched between conditions, that the indoor and outdoor meteorological parameters are reasonably similar, and that data are analyzed by a proper statistical model (Miller *et al.*, 2019). Moreover, a major statistical aspect in study designs that involve repeated measures over time, i.e., longitudinal data analysis, is the modelization of the covariance structure of the data as a function of the process that generates them. Suitable research designs and appropriate statistical techniques are then required. The intention of this investigation was to compare the exercise intensity as measured by HR throughout the same progressive incremental running protocol between tests carried out on track and level-grade motorized treadmills in well-trained athletes.

II. METHODS

A. Participants

The HR responses of twelve healthy, highly trained, and experienced male endurance athletes were analyzed. The participants voluntarily agreed to take part in the study, and consent was obtained from them after a detailed explanation of the purposes and procedures. The research was carried out on the basis of the ethical principles of the Declaration of Helsinki of the World Medical Association (World Medical Association, 2013). The age and basic anthropometric characteristics of the subjects, as well as indirect estimations of maximal oxygen uptake (e$\dot{V}O_{2max}$) obtained from the Cooper test (Cooper, 1968; Saghiv & Sagiv, 2020) are summarized in Table I.

TABLE I: MORPHO-FUNCTIONAL CHARACTERISTICS OF THE PARTICIPANTS

	mean ± SD
Age (years)	24.3±2.7
Weight (kg)	69.3±4.6
Height (m)	1.75±0.06
Body mass index (kg·m^{-2})	22.7±1.2
e$\dot{V}O_{2max}$ (ml·kg^{-1}·min^{-1})	65.4±2.8

SD: standard deviation. e$\dot{V}O_{2max}$: estimated maximal oxygen uptake.

B. Study Design

A cross-over longitudinal design was implemented to study the exercise intensity as measured by HR in a progressive incremental running protocol performed in two test conditions: on a 400 m running track, and on a level-grade motorized treadmill. The running protocol had initial and final speeds of 11 km·h^{-1} and 18 km·h^{-1}, and increments of 0.5 km·h^{-1} every 200 m (15 speed stages), covering a distance of 3 kilometers in 12 minutes and 42 seconds. The range of speeds was determined with the purpose of encompassing a reasonably wide extent of typical training and testing intensities for competitive athletes. Short-wave telemetry devices were used to assess HR (Polar Accurex NV; Polar Electro Oy, Kempele, Finland). Operational reasons did not allow for randomizing the temporal order of the tests in track and treadmill. The subjects suspended high-intensity activities during the testing period. Prior to both evaluations, the participants completed a specific warm-up routine.

C. Track Test

The track tests were performed on an outdoor 400 m running track. With the aim of controlling the running speed, checkpoints were located every one hundred meters along the track, where test assistants with a running pace chart and a digital chronometer were positioned. Two extra test assistants were positioned in two of these checkpoints (separated by two hundred meters from each other) to collect the HR data, which were orally informed by the runners. The subjects were evaluated in pairs (two hundred meters apart from each other), and the twelve tests were completed in one 2-hour morning session. An identical habituation trial was performed seven days before, with the purpose of familiarizing the participants with the task; they were further instructed to adjust softly the running speed when required, avoiding abrupt accelerations and decelerations. Stable meteorological conditions were observed during the testing session, with sunny weather and nearly no wind (daily average values: temperature = 20.4 °C, relative humidity = 58%, atmospheric pressure = 1015.8 hPa, and wind speed = 13.3 km·h^{-1}).

D. Treadmill Test

The treadmill tests were carried out under laboratory conditions, with negligible variations in the room temperature (around 21 °C). The running protocol was performed using a level-grade motorized treadmill (Technogym Run Race HC 1200; Technogym SpA, Gambettola, Italy). In order to manage the increments in the treadmill belt speed, collect the HR data, and supervise the testing protocol, a total of four qualified personnel took part in the evaluation process. The twelve tests were completed six days after the track tests, in two successive 2-hour morning sessions.

E. Statistical Analysis

Data of age, basic anthropometric characteristics, and $e\dot{V}O_{2max}$ of the participants were summarized as mean ± standard deviation (SD). A line chart was produced to illustrate the HR behavior along the running tests on the track and treadmill. The HR data were analyzed by means of a linear mixed-effects model, with Test (track and treadmill) and Speed stage (1st to 15th) and their interaction as fixed factors, and the Subject as a random factor. A suitable correlation structure was specified for the observations of the same subject within the same test. In order to take into account, the unevenly spaced time points of data collection, an exponential spatial correlation structure was used (Pinheiro & Bates, 2000; Verbeke & Molenberghs, 2009). The variance components were estimated with the restricted maximum likelihood method (REML). The log-likelihood ratio statistic was also calculated to assess the goodness of fit. Statistical significance was set at the 0.05 probability level. The R software environment version 4.0.4 was employed for the analyses (R Core Team, 2021).

III. Results

The HR difference between track and treadmill running was not statistically significant ($F = 0.06$, $p = 0.81$). The mean HR response (± standard error) averaged over the fifteen speed stages was 165.2±2.4 beats·min^{-1} on track, and 165.5±2.4 beats·min^{-1} on the treadmill. Moreover, the interaction effect between the factors Test and Speed stage was also not statistically significant ($F = 1.32$, $p = 0.19$). It was observed that the mean differences between track and treadmill across the fifteen speed stages were, in absolute values, less than 2 beats·min^{-1} (from −1.8 to 1.6 beats·min^{-1}). The mean HR responses with their corresponding 95% confidence intervals (95% CI) in all the speed stages for the track and treadmill are reported in Table II. The average HR behavior throughout the running protocol in both test conditions is depicted in Fig. 1.

TABLE II: Mean Heart Rate (95% CI) in Track and Treadmill in the Fifteen Speed Stages

Speed	Track	Treadmill
11.0	136.6 (131.5 to 141.7)	138.0 (132.9 to 143.1)
11.5	145.1 (140.0 to 150.2)	144.9 (139.8 to 150.0)
12.0	147.7 (142.6 to 152.7)	147.8 (142.7 to 152.8)
12.5	153.3 (148.3 to 158.4)	152.1 (147.0 to 157.2)
13.0	156.6 (151.5 to 161.7)	158.3 (153.2 to 163.3)
13.5	161.1 (156.0 to 166.2)	161.7 (156.6 to 166.7)
14.0	163.1 (158.0 to 168.2)	164.4 (159.3 to 169.5)
14.5	167.3 (162.2 to 172.3)	167.7 (162.6 to 172.7)
15.0	171.0 (165.9 to 176.1)	172.1 (167.0 to 177.2)
15.5	172.8 (167.8 to 177.9)	172.3 (167.3 to 177.4)
16.0	174.3 (169.2 to 179.3)	176.1 (171.0 to 181.2)
16.5	177.5 (172.4 to 182.6)	178.8 (173.7 to 183.8)
17.0	180.6 (175.5 to 185.7)	179.9 (174.8 to 185.0)
17.5	184.3 (179.3 to 189.4)	183.0 (177.9 to 188.1)
18.0	186.7 (181.6 to 191.7)	185.1 (180.0 to 190.2)

Heart rate values expressed in beats·min^{-1}. Speed expressed in km·h^{-1}.

The exponential spatial correlation structure used to model the covariation between the observations of the same subject within the same test provided a significantly better fit than the independent structure (log-likelihood ratio statistic = 216.06, $p < 0.0001$). The variance estimate of the random effect (Subject) was also of statistical significance (59.99, $p = 0.01$).

European Journal of Sport Sciences
www.ej-sport.org

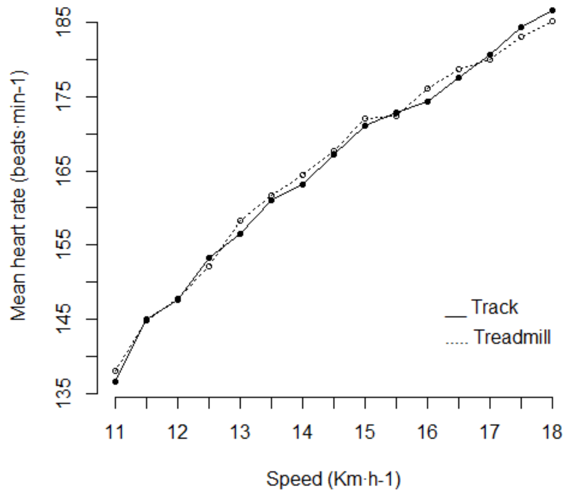

Fig. 1. Heart rate behavior in track and treadmill across the running speeds.

IV. Discussion

Many maximal and submaximal running-based protocols of exercise testing have been designed to be carried out in the field and on the treadmill under laboratory conditions. The exercise intensity is often supposed to be equivalent in both situations, and HR is a conventional and extensively used measure to asses it (American College of Sports Medicine, 2014; Gibson *et al.*, 2019; Kenney *et al.*, 2019; Miller *et al.*, 2019).

Different authors have analyzed comparatively the HR responses obtained when running overground and on a level-grade treadmill. Some of them proposed constant speed protocols, and others implemented variable speed protocols. On one hand, some studies found no significant differences. And on the other hand, other works reported significant differences; however, the direction of these differences was not the same in all cases.

Initially, Brookes *et al*. (1971) contrasted the HR values gathered from three male runners who run overground (road) and on a level-grade motorized treadmill at four speeds (9.6, 12.9, 16.1, and 19.3 $km \cdot h^{-1}$), which resulted in non-statistically significant differences. Also, Maksud *et al*. (1971) analyzed the HR changes that fifteen male college students showed when running on a 140-yard banked indoor track and on a level-grade motorized treadmill at three speeds (11.3, 16.1 and 19.3 $km \cdot h^{-1}$). Sample differences of no statistical significance were found between the two experimental conditions. Afterward, McMurray *et al*. (1988) and White *et al*. (1998) compared the HR responses of trained runners (8 males, and 12 males and 6 females, respectively) during overground (track) and level-grade motorized treadmill running using constant workload protocols (16.1 km at 70% $\dot{V}O_{2max}$, and 60 min at 75% $\dot{V}O_{2max}$, respectively). The first of these works found no significant results, and significantly lower HR changes were registered in the overground running in the second work ($p < 0.05$). As well, Yngve *et al*. (2003) (n = 28; 14 males and 14 females) and Chu *et al*. (2010) (n = 26; all males) assessed the HR responses of young adults by means of constant speed protocols, which consisted of 5 min run and 20 min run, respectively, in both cases at individually selected speeds. Yngve *et al*. (2003) did not find statistically significant differences between overground (indoor track) and level-grade motorized treadmill. On the contrary, Chu *et al*. (2010) detected significantly higher HR values ($p < 0.05$) in overground running (outdoor track).

Alternatively, Ceci and Hassmén (1991) measured the physiological strain produced during running on a flat motorized treadmill and an outdoor 500 m track in eleven physically active men aged between 33 and 65 years; two tests sessions were performed separated by a mean time of 27 days. The two running conditions were compared at three subjective intensity levels based on the Rating of perceived exertion (Borg, 1962; Borg, 1970; Robertson & Noble, 1997). Among other findings, they reported higher HR values for track in both test sessions at the three levels of exercise ($p < 0.001$). On the other hand, in an attempt to establish the treadmill gradient that better reflects the energy cost of outdoor running, Jones and Doust (1996) found no significant differences when comparing the HR measurements collected in motorized treadmill tests at 0% (flat) and 1% grades with the ones obtained in a level road test. A series of six running speeds were considered, and the subjects under study were nine trained male runners (24.9±5.2 years); the range of speeds was from 2.92 to 5.00 $m \cdot s^{-1}$ (10.5 to 18.0 $km \cdot h^{-1}$), which was almost identical to the one proposed in the present study. As well, using a similar incremental running protocol, Kunduracioglu *et al.*

European Journal of Sport Sciences
www.ej-sport.org

(2007) examined the HR responses that twenty-two young male football players (17.9±0.8 years) showed on an artificial turf 120 m hexagonal track and a zero-grade motorized treadmill. The players revealed significantly higher HR values in track ($p < 0.001$) at speeds equal to or greater than 12 km·h^{-1}. In addition, no statistically significant results were found when assessing the HR differences between track and treadmill at a fixed blood lactate concentration of 4 mmol·L^{-1}. Conversely, Di Michele *et al*. (2009) found significantly lower HR values ($p < 0.05$) in a multistage running test performed on a 250 m oval circuit that was traced upon a synthetic turf pitch, in comparison to the same test carried out on a motorized treadmill at level grade and under laboratory conditions, evaluated at three lactate thresholds (2, 3 and 4 mmol·L^{-1}). As well, the HR values in these two running conditions showed no significant differences when compared with the ones obtained by an identical test carried out on a 250 m oval circuit traced upon natural grass. Significant differences were also detected at the speed stages (8, 10, 12, and 14 km·h^{-1}), but only between the synthetic turf and natural grass conditions ($p < 0.05$). The subjects examined were eighteen young male football players (17.4±0.8 years). And more latterly, Köklü *et al*. (2020) also analyzed the HR responses of male football players (n = 19; 17.6±1.1 years) by means of a two-part incremental running test, with speeds of 8, 10, 12, and 14 km·h^{-1} in the first part, and from 15 km·h^{-1} and with increments of 1 km·h^{-1} until volitional exhaustion in the second part. Three experimental conditions were compared: motorized treadmill running under laboratory conditions, field running with 180-degree change of direction every 100 meters, and field running on a circular 100 m track. The field tests were performed on an artificial grass pitch. Significantly higher HR values were observed in the field tests only at the 12 km·h^{-1} stage ($p < 0.05$).

With a different approach, a systematic review published by Miller *et al*. (2019) compiled thirty-four cross-over studies carried out on adults where physiological, perceptual, and performance data collected during overground and treadmill running were compared. As part of the same work, they as well presented the results of a meta-analysis based on those studies. Pooled results indicated that during submaximal running there was a reduction, though not statistically significant, of the heart rate response on level-grade motorized treadmill compared to overground (3 beats·min^{-1}, p = 0.546). Pooled results also showed that near-maximal treadmill running (≥ 80% $\dot{V}O_{2max}$) at 0% grade reduced significantly the heart rate response with respect to overground running (3 beats·min^{-1}, p = 0.011). And at maximal intensity, the difference between the two running conditions was even lower and of no statistical significance (1 beat·min^{-1}, p = 0.518).

On the other hand, potential non-equivalences in energy cost between treadmill and overground locomotion may likely have a noteworthy effect on the cardiac response. The comparison between these two running conditions has also been approached in the literature from a mechanical point of view. To this respect, Van Ingen Schenau (1980) reviewed many research papers about differences in kinematics or in energy consumption. Using theoretical energy calculations, he showed that, as long as the treadmill supplies a stable speed, there are no such mechanical differences. And more recently, Van Hooren *et al*. (2020) carried out a systematic review and meta-analysis based on thirty-three cross-over studies, concluding that "Spatiotemporal, kinematic, kinetic, muscle activity, and muscle-tendon outcome measures are largely comparable between motorized treadmill and overground running." (p. 786).

The standardized environment of the laboratory treadmill tests, which are not altered by meteorological or terrain conditions, makes them ideal for the assessment of variables related to exercise intensity. However, despite some confounding factors that may affect the outcomes, such as the ambient temperature or the wind speed, field tests are also employed; a standardized environment is frequently tacitly assumed. In addition, specific kinematic and dynamic aspects of field and treadmill running that might have influence are also often discarded. In the present study, there was not found a statistically significant difference between the mean HR responses obtained using the same running protocol in an outdoor 400 m track and in a level-grade motorized treadmill. Such difference (averaged over the fifteen speed stages) was less than 0.3 beats·min^{-1} (p = 0.81). Furthermore, the level of exercise intensity seemed not to affect the magnitude of this difference, which throughout all the speed stages showed a fluctuation of less than 2 beats·min^{-1}. Said in more formal terms, the interaction effect between the Test and Speed stage was of no statistical significance (p = 0.19). It is worthwhile to mention that the absence of statistically significant HR differences between the overground and level-grade treadmill conditions was the result most frequently found in the reviewed literature. Special care was taken to minimize different forms of possible bias in the study design. The environmental bias was limited by the fact that the indoor and outdoor meteorological parameters were reasonably comparable in the track and treadmill tests. The intensity bias was straightforwardly restricted; the running speed stages and the distance covered were matched by means of using identical protocols in both test conditions. And in order to avoid statistical bias as far as possible, a suitable statistical model was adjusted, using modern techniques to accurately portray the data-generating process.

The type of design implemented in this research enabled us to assess comparatively the heart rate responses observed during running on a track and level-grade treadmill for more than twelve minutes and throughout fifteen increasing levels of exercise intensity. As well, estimates of physical workload by using heart rate are provided as a function of running speed, which becomes valuable information and offer key guidance for athletes, trainers, and health and sports professionals in general involved in aspects related to cardiorespiratory endurance. On the other hand, the longitudinal nature of the data entailed a scenario of correlation between the subsequent observations of the same subject over time, which is also referred to as intra-individual correlation (Liu, 2015). The statistical model applied captured this correlation structure, as well as control for the covariation between measures of the same subject in the two test conditions (Pinheiro & Bates, 2000). Ignoring these covariations represents a violation of the assumption of independence between observations, and may lead to spurious inference results (Weber & Skillings, 2000). Nevertheless, a longitudinal data structure with shorter time or space intervals between measurements may possibly provide more accurate and trustworthy results. It is also worth mentioning that although the subjects under study were highly trained and experienced runners, the strategy used in the track test to adjust the established running speeds is a limitation of this research. Technological devices such as the GPS (Global Positioning System), an audio signal player, or a system of flashing lights may be more accurate options. However, a larger number of checkpoints could also be an overcoming alternative. Furthermore, in order to avoid possible sequential effects in the track and treadmill tests as far as possible, a randomized testing order is also recommended. Therefore, future research would be advisable to provide further scientific evidence.

V. Conclusion

Exercise intensity as measured by heart rate showed comparable mean responses in track and level-grade treadmill running across a wide range of speeds in well-trained athletes. A non-significant interaction effect between the factors Test and Speed stage was observed; the contrast between running on track and treadmill did not appear to depend on the level of physical workload, at least in the range of speeds studied. Nonetheless, more frequent heart rate monitoring during the tests may yield more robust conclusions.

Acknowledgment

The authors are especially thankful to Néstor A. Lentini and Gerónimo M. Gris for their scientific support of this work.

Conflict of Interest

The authors declare that they do not have any conflict of interest.

References

American College of Sports Medicine. (2014). *Guidelines for graded exercise testing and exercise prescription*. 9th ed. Philadelphia, PA: Lippincott Williams & Wilkins.

Borg, G. A. V. (1962). *Physical performance and perceived exertion*. Lund, Sweeden: Gleerup.

Borg, G. A. V. (1970). Perceived exertion as an indicator of somatic stress. *Scand J Rehabil Med*, 2-3, 92-98.

Brookes, F. B., Knibbs, A. A., Pantlin, C. M., & Wilson, J. K. (1971). An investigation into the biomechanical and physiological differences between road and treadmill running. *Res Pap Phys Educ*, 2(2), 28-35.

Ceci, R., & Hassmén, P. (1991). Self-monitored exercise at three different RPE intensities in treadmill vs field running. *Med Sci Sports Exerc*, 23(6), 732-738. https://doi.org/10.1249/00005768-199106000-00013.

Chu, C. Y., Lu, S. Y., & Lin, K. F. (2010). Influences of exercise experience and exercise settings on heart rate responses during self-selected intensity exercises. *J Exerc Sci Fit*, 8(2), 73-77. https://doi.org/10.1016/S1728-869X(10)60011-1.

Cooper, K. H. (1968). A means of assessing maximal oxygen intake. Correlation between field and treadmill testing. *J Amer Med Assoc*, 203(3), 201-204. DOI: 10.1001/jama.1968.03140030033008.

Di Michele, R., Di Renzo, A. M., Ammazzalorso, S., & Merni, F. (2009). Comparison of physiological responses to an incremental running test on treadmill, natural grass, and synthetic turf in young soccer players. *J Strength Cond Res*, 23(3), 939-945. DOI: 10.1519/JSC.0b013e3181a07b6e.

Gibson, A. L., Wagner, D., & Heyward, V. (2019). *Advanced fitness assessment and exercise prescription*. 8th ed. Champaign, IL: Human Kinetics.

Jones, A. M., & Doust, J. H. (1996). A 1% treadmill grade most accurately reflects the energetic cost of outdoor running. *J Sports Sci*, 14(4), 321-327. https://doi.org/10.1080/02640419608727717.

Kenney, L., Wilmore, J. H., & Costill, D. L. (2019). *Physiology of sport and exercise*. 7th ed. Champaign, IL: Human Kinetics.

Köklü, Y., Alemdaroğlu, U., Demirhan, R., & Arslan, Y. (2020). A comparison of incremental running field and treadmill tests in young soccer players. *J Hum Kinet*, 73, 193-201. https://doi.org/10.2478/hukin-2019-0143.

Kunduracioglu, B., Guner, R., Ulkar, B., & Erdogan, A. (2007). Can heart rate values obtained from laboratory and field lactate tests be used interchangeably to prescribe exercise intensity for soccer players? *Adv Ther*, 24(4), 890-902. https://doi.org/10.1007/BF02849982.

Liu, X. (2015). *Methods and applications of longitudinal data analysis*. New York, NY: Academic Press.

Maksud, M. G., Coutts, K. D., & Hamilton, L. H. (1971). Time course of heart rate, ventilation, and $\dot{V}O_2$ during laboratory and field exercise. *J Appl Physiol*, 30(4), 536-539. https://doi.org/10.1152/jappl.1971.30.4.536.

McMurray, R. G., Berry, M. J., Vann, R. T., Hardy, C. J., & Sheps, D. S. (1988). The effect of running in an outdoor environment on plasma beta endorphins. *Ann Sports Med*, 3(4), 230-233.

Miller, J. R., Van Hooren, B., Bishop, C., Buckley, J. D., Willy, R. W., & Fuller, J.T. (2019). A systematic review and meta-analysis of crossover studies comparing physiological, perceptual and performance measures between treadmill and overground running. *Sports Med*, 49(5), 763-782. https://doi.org/10.1007/s40279-019-01087-9.

Montgomery, D. C. (2013). *Design and analysis of experiments*. 8th ed. Hoboken, NJ: John Wiley & Sons, Inc.

Pinheiro, J., & Bates, D. M. (2000). *Mixed-effects models in S and S-PLUS*. New York, NY: Springer; p. 3.

R Core Team. (2021). *R: A language and environment for statistical computing*. R Foundation for Statistical Computing, Vienna, Austria. Available at: http://www.R-project.org/.

Reuter, B. H., & Dawes, J. J. (2016). Program design and technique for aerobic endurance training. In G. G. Haff, & N. T. Triplett (Eds.), *Essentials of strength training and conditioning* (4th ed., pp. 559-582). Champaign, IL: Human Kinetics.

Robertson, R. J., & Noble, B. N. (1997). *Perception of physical exertion: Methods, mediators, and applications. Exercise and sport sciences reviews. Vol. 25*. Baltimore, MD: Williams & Wilkins; pp. 407-452.

Saghiv, M. S., & Sagiv, M. S. (2020). *Basic exercise physiology: Clinical and laboratory perspectives*. Cham, Switzerland: Springer Nature Switzerland AG; p. 171.

SAS Institute Inc. (1999). *SAS/STAT® 9.3 User's guide, Version 8*. Cary, NC: SAS Institute Inc. Available at: http://support.sas.com/documentation/onlinedoc/stat/930/.

Van Belle, G., Fisher, L. D., Heagerty, P. J., & Lumley, T. (2004). *Biostatistics: A methodology for the health sciences*. 2nd ed. Hoboken, NJ: John Wiley & Sons, Inc.

Van Hooren, B., Fuller, J. T., Buckley, J. D., Miller, J. R., Sewell, K., Rao, G., *et al*. (2020). Is motorized treadmill running biomechanically comparable to overground running? A systematic review and meta-analysis of cross-over studies. *Sports Med*, 50(4), 785-813. DOI: 10.1007/s40279-019-01237-z.

Van Ingen Schenau, G, J. (1980). Some fundamental aspects of the biomechanics of overground versus treadmill locomotion. *Med Sci Sports Exerc*, 12(4), 257-261. https://doi.org/10.1249/00005768-198024000-00005.

Verbeke, G., & Molenberghs, G. (2009). Repeated measures and multilevel modeling. In: S. R. Wilson, & C. Burden (Eds.), *Biometrics. Vol. I. Encyclopedia of life support systems* (pp. 237-255). Oxford, United Kingdom: Eolss Publishers Co. Ltd.

Weber, D. C., & Skillings, J. H. (2000). *A first course in the design of experiments: A linear models approach*. Boca Raton, FL: CRC Press LL; pp. 225-226.

White, J. A., Pomfret, D. K., Rennie, S., Wong, J., & Ford, M. (1998). Fluid replacement needs of well-trained male and female athletes during indoor and outdoor steady state running. *J Sci Med Sport*, 1(3), 131-142. https://doi.org/10.1016/S1440-2440(98)80009-1.

World Medical Association (WMA). (2013). *WMA Declaration of Helsinki - Ethical principles for human medical research*. Available at: https://www.wma.net/policies-post/wma-declaration-of-helsinki-ethical-principles-for-medical-research-involving-human-subjects/.

Yngve, A., Nilsson, A., Sjöström, M., & Ekelund, U. (2003). Effect of monitor placement and of activity setting on the MTI accelerometer output. *Med Sci Sports Exerc*, 35(2), 320-326. DOI: 10.1249/01.MSS.0000048829.75758.A0.