TITLE

Test-retest reliability of running economy and metabolic and cardiorespiratory parameters during a multi-stage incremental treadmill test in male middle- and long-distance runners

AUTHORS

Aidan J. Brady[1,2], Mark Roantree[2], Brendan Egan[1,3]

AFFILIATIONS

[1]School of Health and Human Performance, Dublin City University, Dublin, Ireland.
[2]Insight Centre for Data Analytics, Dublin City University, Dublin, Ireland.
[3]Florida Institute for Human and Machine Cognition, Pensacola FL, United States

RUNNING TITLE

Reliability of incremental running test

ORCID ID

Aidan Brady 0000-0002-9427-5771
Mark Roantree 0000-0002-1329-2570
Brendan Egan 0000-0001-8327-9016

CORRESPONDING AUTHOR

Brendan Egan, PhD
School of Health and Human Performance
Dublin City University
Glasnevin, Dublin 9
Ireland
e: brendan.egan@dcu.ie
t: +353 1 7008803
f: +353 1 7008888

ACKNOWLEDGEMENTS

No conflict of interest, financial or otherwise, is declared by the author(s). The governing body providing funding to AB and MR have not reviewed or provided input to the manuscript. **Author contributions:** Conceptualization, methodology, investigation, data curation, formal analysis, writing – original draft: AB. Formal analysis, funding acquisition, supervision, writing – review and editing: MR. Conceptualization, methodology, project administration, supervision, writing – review and editing: BE. **Author funding:** AB and MR are supported by Science Foundation Ireland (SFI) under Grant Number SFI/12/RC/2289_P2, co-funded by the European Regional Development Fund.

ABSTRACT WORD COUNT:	**249**
WORD COUNT:	**4095**
REFERENCES:	**39**
FIGURES	**1**
TABLES	**3**

ABSTRACT

This study investigated the test-retest reliability of running economy (RE) and metabolic and cardiorespiratory parameters related to endurance running performance using a multi-stage incremental treadmill test. On two occasions separated by 21- to 28-d, twelve male middle- and long-distance runners ran at 10, 11, 12, 13, and 14 km·h^{-1} for 8 min each stage, immediately followed by a ramp test to volitional exhaustion. Carbohydrate (10% maltodextrin solution) was consumed before and during the test to provide ~1 g·min^{-1} of exercise. RE, minute ventilation ($\dot{V}E$), oxygen consumption ($\dot{V}O_2$), carbon dioxide production ($\dot{V}CO_2$), respiratory exchange ratio (RER), heart rate (HR), ratings of perceived exertion (RPE), and blood glucose and lactate concentrations, were recorded for each stage and at volitional exhaustion. Time to exhaustion (TTE) and peak oxygen consumption ($\dot{V}O_{2peak}$) during the ramp test were also recorded. Absolute reliability, calculated as the coefficient of variation (CV) between repeated measures, ranged from 2.3 to 3.1% for RE, whereas relative reliability, calculated as the intraclass correlation coefficient (ICC), ranged from 0.42 to 0.79. $\dot{V}E$, $\dot{V}O_2$, $\dot{V}O_{2peak}$, $\dot{V}CO_2$, RER, and HR had CVs of 1.1 to 4.3% across all stages. TTE and RPE had CVs of 7.2% and 2.3 to 10.8%, respectively, while glucose and lactate had CVs of 4.0 to 17.8%. All parameters, except for blood glucose, were demonstrated to have *good-to-excellent* relative reliability assessed by ICC. Measures of RE, $\dot{V}O_{2peak}$, and TTE were reliable during this two-phase multi-stage incremental treadmill test in a cohort of trained and highly-trained male middle- and long-distance runners.

Keywords: lactate threshold; maximal oxygen consumption; running economy; time to exhaustion; time trial.

INTRODUCTION

Improvements in physiological and performance indices such as running economy (RE), blood lactate, maximal oxygen consumption ($\dot{V}O_{2max}$), and time to exhaustion (TTE) are common targets of nutrition and training interventions. These indices are strong determinants of middle- and long-distance running performance, and useful in the monitoring and prescription of training (Barnes & Kilding, 2015a; Jamnick et al., 2018; Jones et al., 2021). Determination of these indices in unison typically requires a laboratory-based, multi-stage incremental exercise protocol, although no standardized protocol is currently employed (Barnes & Kilding, 2015a; Jamnick et al., 2018). A diverse range of protocols are consequently available, yet only a limited number of studies have adequately investigated the test-retest reliability of relevant physiological and performance indices (Blagrove et al., 2017; Saunders et al., 2004; Shaw et al., 2013). Rather, these physiological and performance indices have often been examined separately and at a single intensity (Brisswalter & Legros, 1994; Morgan et al., 1991; Pereira & Freedson, 1997), and therefore have limited utility in research or practice.

RE is recognized as an important discriminator of performance between athletes of a similar training status, with small improvements in RE capable of eliciting meaningful improvements in competitive long-distance running performance (Barnes & Kilding, 2015a). Consequently, a wide range of training (e.g., resistance exercise, plyometrics) and nutrition (e.g., caffeine, dietary nitrate) strategies specifically targeting RE have been investigated (Barnes & Kilding, 2015b). However, changes in RE in response to an intervention must be interpreted alongside the measurement error associated with the exercise protocol, and the inherent intra-individual variation in the outcome measure (Saunders et al., 2004). If the measurement error is greater than the apparent improvement, then the effect is unlikely to constitute a 'true' change (Atkinson & Nevill, 1998; Hopkins, 2000). Therefore, establishing test-retest reliability is central to interpreting the efficacy of a given intervention and providing objective data on performance indicators (Atkinson & Nevill, 1998; Burke, 2017; Currell & Jeukendrup, 2008).

Protocols facilitating the simultaneous, and reliable measurement of TTE and/or peak oxygen consumption ($\dot{V}O_{2peak}$) may be useful to coaches and practitioners as additional indices of performance (Bentley et al., 2007). Yet, the exercise protocol design, particularly the stage duration, can influence the validity and reliability of RE, blood lactate, and $\dot{V}O_{2max}$ (Bentley et al., 2007; Jamnick et al., 2018). During continuous exercise, oxygen consumption ($\dot{V}O_2$) and heart rate (HR) can attain a steady state in 3–5 min while longer durations of ≥6 min have been recommended for the valid measurement of blood lactate, particularly with non-elite participants (Bentley et al., 2001; Foxdal et al., 1996; Kuipers et al., 2003). For example, in a cohort of active males and females, the measurement error was lower across a range of blood lactate concentrations, including lactate threshold, when using 8 min stage durations compared to 4 min, and also with a lower average bias and narrower limits of agreement (Gavin et al., 2014).

Incremental running protocols involving measurements of blood parameters typically utilize a discontinuous approach, providing a safer, more efficient method of sample collection (Gullstrand et al., 1994). During prolonged exercise trials, especially those with a nutrition intervention requiring bolus ingestion during the interval between stages, longer durations than the recommended 30 s interval (Gullstrand et al., 1994) are likely required to allow for ingestion in a timely manner alongside the concurrent measurement of metabolic and cardiorespiratory parameters. Many studies examining the reliability of RE and blood lactate concentrations have done so in the absence of carbohydrate ingestion (Blagrove et al., 2017; Saunders et al., 2004; Shaw et al., 2013). Because endurance athletes routinely consume carbohydrate during training and competition (Burke et al., 2019), it is arguable that carbohydrate ingestion should be included during laboratory-based testing, particularly as these measures are used to tailor training programs and develop pacing strategies.

The aim of this study was to investigate the test-retest reliability of RE, blood glucose and blood lactate concentration, minute ventilation ($\dot{V}E$), $\dot{V}O_2$, carbon dioxide production ($\dot{V}CO_2$), respiratory exchange ratio (RER), HR, ratings of perceived exertion (RPE), and TTE measured during submaximal and maximal exercise using a two-phase multi-stage incremental treadmill test performed by trained middle- and long-distance runners on two separate occasions under standardized conditions that included carbohydrate ingestion.

METHODS

Participants

Twelve trained male middle- and long-distance runners [mean±standard deviation (SD); age, 37.0±6.6 y; height, 1.82±0.06 m; body mass, 75.1±6.6 kg; sum of 8 skinfolds, 74.7±18.5 mm] provided written informed consent to participate after written and verbal explanation of the study procedures. Ethical approval was obtained from the ANONYMISED Research Ethics Committee (permit: 2021_ANONYMISED) in accordance with the Declaration of Helsinki, with the exception of registration in a trial database. All participants had a background in endurance running, were training at least 4 times per week and/or for at least 4 hours per week, and were classified as Tier 2/Trained and Tier 3/Highly Trained level athletes (McKay et al., 2022). Participants were also free from injury for the previous 6 months, and familiar with treadmill running.

Study design

Participants visited the human performance laboratory at ANONYMISED on two separate occasions separated by 21- or 28-d between March and December of 2022. Each visit involved the same exercise protocol, consisting of submaximal treadmill running at five different speeds during which several metabolic and cardiorespiratory parameters were assessed, followed by a ramp test to volitional exhaustion to assess TTE and $\dot{V}O_{2peak}$. Participants also consumed a carbohydrate solution prior to commencing exercise, and between the second and third submaximal running speeds during each visit. Test-retest reliability for RE, blood glucose and blood lactate concentration, $\dot{V}E$, $\dot{V}O_2$, $\dot{V}CO_2$, RER, HR, RPE, and TTE was examined using the coefficient of variation (CV) and intraclass correlation coefficient (ICC). The primary outcome measure was reliability of RE at each submaximal running speed. Secondary outcome measures were reliability of blood glucose and blood lactate concentration, $\dot{V}E$, $\dot{V}O_2$, $\dot{V}CO_2$, RER, HR, RPE, and TTE.

Pre-test preparation

Each visit to the laboratory took place on the same day of the week and at the same time of day ±1 h. All exercise tests commenced between 1800 and 2200 h. Participants were asked to abstain from alcohol for 48 h prior, caffeine 5 h prior, and refrain from strenuous exercise training for 24 h prior to each visit. Participants were also asked to consume a standardized meal plan designed to provide 30 $kcal \cdot kg^{-1}$ with a macronutrient composition of 60% carbohydrate, 20% protein, and 20% fat for the day before and day of each visit, with the last meal being consumed 3 h prior to arrival to the laboratory. Compliance with these standardization measures was assessed verbally upon arrival for each visit. Hydration status was assessed from a urine sample provided upon arrival to the laboratory by using a digital handheld refractometer (PEN-Urine S. G., ATAGO CO.,LTD., Tokyo, Japan) to measure urine specific gravity. Participants were instructed to continue with their habitual dietary and training routines between visits, and provide their activity data for the 7-d period prior to their first visit, and each subsequent week of participation thereafter, from their GPS-enabled smart watch (Garmin, Kansas, USA) for verification. The mean±SD for ambient temperature, relative humidity, and barometric pressure was 19.6±0.4°C, 56.1±9.8%, and 756.7±6.2 mmHg, respectively, during the first visit, and 19.8±0.5°C, 53.4±5.9%, and 755.5±8.4 mmHg, respectively, during the second visit (all $P>0.05$).

Anthropometry

Body mass was measured to the nearest 0.1 kg using a calibrated digital scale (model 704, SECA, Hamburg, Germany). Height was measured to the nearest 0.01 m using a wall-mounted stadiometer (model 200, SECA, Hamburg, Germany). Subcutaneous fat tissue was measured to the nearest 0.1 mm by an International Society for the Advancement of Kinanthropometry certified practitioner from 8 anatomical locations (triceps, subscapular, biceps, iliac crest, supraspinale, abdomen, front thigh, and medial calf) using a Harpenden skinfold caliper (Baty International, West Sussex, UK).

Protocol design

A two-phase multi-stage incremental running protocol performed on a motorized treadmill (T170; COSMED, Rome, Italy) was used during each visit. The first phase involved five stages at a 1% gradient, beginning at 10 $km{\cdot}h^{-1}$ and increasing in increments of 1 $km{\cdot}h^{-1}$. Each stage was 8 min in duration interspersed with a 30 s rest interval, except for after the second stage when the rest interval was increased to 2 min to allow participants sufficient time to ingest a second bolus of a carbohydrate solution. The solution provided consisted of 10% carbohydrate with maltodextrin (100% Maltodextrin Carbs, Myprotein, UK) as the carbohydrate source and flavored with a lemon sweetener (FlavDrops™, Myprotein, UK). The solution provided carbohydrate at a rate of ~1.0 g·min^{-1} of exercise. Participants consumed 500 mL 20 min prior to exercise and 300 mL after completion of the second submaximal running stage (11 km·h^{-1}). The second phase involved a ramp test to volitional exhaustion, commencing immediately after the rest period of the final 8 min stage (14 km·h^{-1}). During the ramp test, stage duration was reduced to 1 min with the running speed increasing by 1 km·h^{-1} every minute up to 16 km·h^{-1}, after which treadmill gradient was increased by 1% until volitional exhaustion.

To minimize the potential for measurement error due to features of the exercise protocol, the present study followed established criteria for the measurement of RE and blood lactate where possible, including the stage duration (Bentley et al., 2007; Foxdal et al., 1996), gradient (Jones & Doust, 1996), and rest interval (Gullstrand et al., 1994). A stage duration ≥6 min is likely longer than that required for $\dot{V}O_2$ and HR to attain a steady-state (Burnley & Jones, 2007; Foxdal et al., 1996), but this duration is necessary for accurate assessment of the blood lactate response to incremental exercise (Bentley et al., 2001; Foxdal et al., 1996; Kuipers et al., 2003). Longer stages are therefore, appropriate when testing endurance athletes given the greater relation of RE and blood lactate with successful performance compared to $\dot{V}O_{2max}$ (Barnes & Kilding, 2015a; Jones, 2006; Saunders et al., 2004; Sjödin & Svedenhag, 1985).

Capillary blood samples of 0.6 µL and 0.3 µL were taken during the rest period between each stage from the earlobe for determination of blood glucose (FreeStyle Optium Neo, Abbott Industries, Illinois, USA), and blood lactate concentration (Lactate Pro 2, Arkray, Kyoto, Japan), respectively. Capillary blood samples were also taken prior to the first bolus of carbohydrate, immediately prior to the start of exercise, and immediately upon cessation of exercise.

Expired air was collected (7450 V2 Mask, Hans Rudolph, Inc., Kansas, USA) and analyzed throughout the entire test using an automated breath-by-breath (BxB) system (Quark RMR, COSMED, Rome, Italy). Participants were fitted with a HR monitor (H10, Polar, Kempele, Finland), and wore the same running shoes for both visits. RE, $\dot{V}E$, $\dot{V}O_2$, $\dot{V}CO_2$, and RER were calculated from a 60 s rolling average of BxB measurements during the last 90 s of each 8 min stage. HR was also calculated from a 60 s rolling average of measurements during the last 90 s. Peak values of $\dot{V}E$, $\dot{V}O_2$, $\dot{V}CO_2$, and RER were recorded as the highest 20 s average observed during the ramp test. The total time completed during the ramp test to volitional exhaustion was recorded as the TTE. RPE (Borg 6–20 scale) was recorded during the last 30 s of each stage.

Statistical analysis

Analysis was performed using GraphPad Prism v.9.4 (GraphPad Software Inc, California, USA). Data are presented as mean±SD unless otherwise stated. A one-way analysis of variance was performed to compare the number of training sessions, weekly distance, and the average heart rate (AHR) and maximum heart rate (MHR) of training each week. A paired samples t-test was used to compare the differences between the first and second visit for RE, blood glucose and blood lactate concentration, $\dot{V}E$, $\dot{V}O_2$, $\dot{V}CO_2$, RER, HR, RPE, and TTE. Standardized estimates of effect size were calculated using the mean difference and pooled SD and are presented as the adjusted Hedges' *g*, as recommended for correlated groups (Lakens, 2013). The significance level was set at $\alpha \leq 0.05$ for all tests. The CV was used as a measure of *absolute* reliability for each outcome variable and calculated as the SD of the difference between repeated measurements divided by the mean of the repeated measurements, then expressed as a percentage. A two-way random, single measure ICC was performed for each outcome variable to provide a measure of *relative* reliability with values interpreted as follows: <0.4, "*poor*"; 0.4–0.5, "*fair*"; >0.5–0.75, "*good*"; and >0.75 "*excellent*" (Cicchetti, 1994). Confidence intervals are reported for CV and ICC values as the 95% confidence interval (CI). The smallest worthwhile change (SWC) was calculated as 0.2 x the pooled SD for all metabolic and cardiorespiratory parameters (Hopkins, 2000).

RESULTS

Body mass was 75.1±6.6 kg and 74.8±6.2 kg at the first and second visit, respectively (P=0.217, CV=0.6%, 95% CI=0.2 to 0.9%, ICC=0.99, ES=0.05). Hydration status was 1.009±0.006 $g{\cdot}mL^{-1}$ and 1.011±0.005 $g{\cdot}mL^{-1}$ at the first and second visit, respectively (P=0.481, CV=0.3%, 95% CI=0.1 to 0.5%, ICC=0.37, ES=0.22). There was no significant difference in the number of training sessions (P=0.480), weekly distance (P=0.782), AHR (P=0.788) or MHR (P=0.645) during the study period. Descriptive analysis of all metabolic and cardiorespiratory parameters are presented in Table 1–3 alongside the CV and ICC and their respective 95% CIs, and the SWC.

No differences were found in RE between visits across each submaximal running speed (Table 1; Figure 1). The absolute reliability of RE ranged from 2.3 to 3.1%CV, whereas relative reliability declined as running speed was increased (ICC at 10 $km{\cdot}h^{-1}$ of 0.74 vs. 0.42 at 14 $km{\cdot}h^{-1}$) (Table 1). $\dot{V}E$, $\dot{V}O_2$, $\dot{V}O_{2peak}$, $\dot{V}CO_2$, RER, and HR had CVs between 1.1 to 4.3% across all stages (Table 1; Table 2). TTE and RPE had CVs of 7.2% and 2.3 to 10.8%, respectively (Table 2), while blood glucose and blood lactate had CVs of 4.0 to 17.8% (Table 3). The relative reliability for $\dot{V}E$, $\dot{V}O_2$, $\dot{V}O_{2peak}$, $\dot{V}CO_2$, HR, TTE, and blood lactate was generally *excellent* (ICC>0.75), for RPE and RER was generally *good* (ICC>0.5 to 0.75), and for blood glucose was generally *poor-to-fair* (ICC<0.05) (Table 1; Table 2; Table 3).

DISCUSSION

This study investigated the test-retest reliability of RE and metabolic and cardiorespiratory parameters related to endurance running performance measured during a two-phase multi-stage incremental treadmill test. RE, $\dot{V}E$, $\dot{V}O_2$, $\dot{V}O_{2peak}$, $\dot{V}CO_2$, RER, and HR all demonstrated adequate absolute reliability with each having a CV of <5% across the various running speeds. TTE, RPE, blood glucose and blood lactate were observed to be less reliable with the CV ranging from 4.0 to 17.8%. Relative reliability assessed by ICC demonstrated greater variation across the parameters measured, and was influenced by running speed. For example, the ICC for RE ranged from 0.79 at 11 $km.h^{-1}$ to 0.42 at 14 $km.h^{-1}$. All other parameters, except for blood glucose, were demonstrated to have *good-to-excellent* relative reliability.

The average CV of RE across the five submaximal running stages is in line with the 1–5%CV reported elsewhere (Barnes & Kilding, 2015a). The 3.1%CV for RE at 14 $km{\cdot}h^{-1}$ in the present study is similar to the 2.4%CV previously reported for elite middle- and long-distance runners at the same speed when expressed as $\dot{V}O_2$ in absolute terms (Saunders et al., 2004). Similar values for absolute reliability of RE have been reported in highly- and moderately-trained runners who ran at approximately 85% of their lactate turnpoint (14.4 $km{\cdot}h^{-1}$ and 11.7 $km.h^{-1}$, respectively) for 15 min once weekly for three consecutive weeks (1.8–2.0%CV) (Pereira & Freedson, 1997), in trained runners who ran at 12 $km{\cdot}h^{-1}$ on two occasions separated by two days (1.3%CV) (Morgan et al., 1991), and in trained runners who ran at a range of speeds from 9 $km{\cdot}h^{-1}$ to 13 $km{\cdot}h^{-1}$ for 10 min each five times per week for four consecutive weeks (2.5–3.1%CV) (Williams et al., 1991). Notably, those studies did not involve the ingestion of carbohydrate, so the similar test-retest reliability of RE in the present study suggests that carbohydrate ingestion immediately prior to and/or during exercise does not negatively impact the absolute reliability of RE. The similar CV for RE across studies is also supportive of previous reports that training status does not appear to influence the reliability of RE with CV values being similar between elite and non-elite participants (Barnes & Kilding, 2015a). Laboratory-based treadmill running tests should consider the inclusion of carbohydrate ingestion in line with recommended exercise nutrition guidelines during training and competition (Thomas et al., 2016; Burke et al., 2019) to improve the ecological validity of the results.

The 4 week timeframe in the present study is considerably longer than previous investigations of the reliability of RE (1–7 days) as reviewed elsewhere (Barnes & Kilding, 2015a). These data indicate that when training frequency, duration, and intensity remain consistent, the absolute reliability of RE is stable over a period of up to 4 weeks. This reliability is an important design consideration for studies such as nutrition interventions involving multiple laboratory visits, or short-term training interventions targeting changes in RE. One caveat when interpreting the present findings is that the data for training load were collected for only 7-d prior to the first visit, rather than the equivalent of the 21- to 28-d of training load prior to the second visit. An observation period of 21- to 28-d prior to the first visit, thereby matching the period of observation for both visits, would be a more comprehensive view of training load over 6 to 8 weeks and facilitate a more conclusive assessment of the impact of habitual training on the reliability of RE.

Although the absolute reliability of RE was largely consistent across the range of submaximal running speeds, there was a progressive decline in relative reliability, assessed by the ICC, as running speeds increased. This reduction is likely due to changes in the rank order of the participants between the first and second visit as illustrated by the difference scores for individual participants. At 11 $km{\cdot}h^{-1}$, which had the highest ICC of 0.79 (*excellent*), changes in RE ranged from -7.0 to 12.7 $mL{\cdot}kg^{-1}{\cdot}km^{-1}$ whereas at 14 $km.h^{-1}$, where the ICC was 0.42

(*fair*), the range was -23.4 to 16.4 $mL.kg^{-1}.km^{-1}$. The influence of running speed on relative reliability is therefore an important consideration when monitoring RE over a prolonged period such as response to a training intervention, or comparing RE between participants in a heterogenous group.

The protocol employed in the present study does not provide a gold-standard assessment of $\dot{V}O_{2max}$ given the prolonged nature of the test (Kuipers et al., 2003), and the absence of a verification bout (Schaun, 2017). Despite these limitations, the reliability observed for $\dot{V}O_{2peak}$, HR_{peak}, and TTE suggests that this protocol could provide a suitable alternative for monitoring training status, particularly because $\dot{V}O_{2max}$ is rarely used for training prescription in an endurance-trained cohort (Jamnick et al., 2020). Furthermore, the CV for $\dot{V}O_{2peak}$ and HR_{peak} are similar to that reported for males of varying athletic standards (Doma et al., 2012; Lourenço et al., 2011; Succi et al., 2023). The CV for TTE protocols involving trained individuals regularly exceeds 10% (Currell & Jeukendrup, 2008), so it is notable that the CV reported here is lower at 7.2%. For comparison, trained male runners who ran to exhaustion at both their 1500 m and 5000 m race pace had a CV of 13.2% and 15.1% for TTE, respectively (Laursen et al., 2007). Similarly, the TTE CV reported for trained male runners who ran at a speed equivalent to 110% of the second ventilatory threshold was 9.2% (Doma et al., 2012). An important distinction is that the protocol in the present study involves an incremental bout resembling a standard graded exercise test whereas previous investigations have typically involved exercise at a constant running speed.

Despite the extended stage duration, the absolute reliability observed for blood lactate and blood glucose was low. Blood lactate was also lower by ~0.2 mM at 11 $km{\cdot}h^{-1}$ and 12 $km{\cdot}h^{-1}$ when comparing the second visit to the first, although there were no differences between visits at faster running speeds including at volitional exhaustion. The low reliability is largely a consequence of expressing the results as a %CV, and the training status of the participants, such that the appearance of lactate deviates minimally across the range of submaximal running speeds, and thereby making small absolute changes appear more pronounced as a %CV. These results should be interpreted in a practical context, where small absolute changes (e.g., 0.1–0.2 mM) may be of limited importance, particularly when exercising below the lactate threshold. Additionally, the average difference between visits for the five submaximal stages is lower than the analytical error of the handheld analyzer used to measure blood lactate, which was reported as ~0.4 mM across a range of intensities (Bonaventura et al., 2015). This error exceeds the SWC for blood lactate at all time points, including peak exercise. In practical terms, the precision of the analyzer is not sufficient to reliably detect small changes in blood lactate concentrations across the range of speeds employed. However, the reliability of blood lactate measured at 14 $km{\cdot}h^{-1}$ in the present study is considerably better than the 52%CV reported for elite middle- and long-distance runners at the same speed using a 4 min stage duration (Saunders et al., 2004). Stage duration should not be compromised for practical convenience, given that the absence of a steady-state can introduce measurement error to blood lactate concentrations at a given speed, and reduce the reliability of the measure (Bentley et al., 2001; Foxdal et al., 1996; Gavin et al., 2014; Kuipers et al., 2003).

The brief extension to the duration of the rest period after the second submaximal running stage was to facilitate the consumption of the second bolus of the carbohydrate solution. The specific timing of this longer interval between stages was because this test protocol was developed as a means to reliably assess RE and other related metabolic and cardiorespiratory parameters for a recent study from our group investigating the effect of a peri-exercise nutrition intervention on RE in a cohort of middle- and long-distance runners (Brady & Egan, 2024). To follow recommended guidelines for carbohydrate intake (Thomas et al., 2016; Burke et al., 2019), participants were required to ingest a bolus of ~300 mL of a carbohydrate solution during the exercise bout, which required the extended rest interval between two of the submaximal running speeds. Additionally, to examine if the nutrition intervention impaired high-intensity exercise performance, and if changes in the oxygen cost of exercise were a consequence of a reduced $\dot{V}O_{2peak}$, a two-phase protocol was required to assess TTE and $\dot{V}O_{2peak}$.

Elite middle- and long-distance runners have a lactate threshold that is considerably higher than the submaximal running speeds used in the present study (Jones et al., 2021). The range of submaximal speeds used may therefore, need to be adjusted in line with an athlete's training status, but providing the protocol design and control of extrinsic factors remains constant, similar levels of reliability are likely. Despite no significant differences in weekly training volume during the study involvement, HR was significantly lower by ~5 beats·min^{-1} during the second visit across the range of submaximal running speeds. The lower HR may be a consequence of a modest training response during the 4 week period between tests, or an order effect of repeat testing, but the absence of differences between tests in most other measured parameters makes it difficult to draw solid conclusions. Interestingly, the mean difference in HR of 4 to 6 beats·min^{-1} between visits is similar to the day-to-day-variability of HR, which has been reported as 1 to 6 beats·min^{-1} (Lambert et al., 1998). Regardless, changes in the HR response should arguably be interpreted with caution when using a crossover design under similar test conditions.

These test-retest reliability data suggest that this is a protocol that could be used for similar research designs in the future, particularly those involving multiple laboratory visits and/or the ingestion of a nutrition supplement. The CV and SWC data may also be used by researchers and practitioners to interpret if changes in RE, blood lactate concentration, $\dot{V}O_{2peak}$, TTE, or any of the other measured metabolic and cardiorespiratory parameters in response to a nutrition or training intervention represent a practically-relevant effect.

In summary, the two-phase multi-stage incremental treadmill test in the present study provided a reliable measure of RE, $\dot{V}O_{2peak}$, TTE and related metabolic and cardiorespiratory parameters when performed twice over a 3 to 4 week period. Measures of blood lactate and glucose were less reliable. Careful consideration of absolute and relative change or difference alongside the CV, ICC, and SWC is required to determine if an observed effect is of practical relevance.

REFERENCES

Atkinson, G., & Nevill, A. M. (1998). Statistical Methods For Assessing Measurement Error (Reliability) in Variables Relevant to Sports Medicine. *Sports Medicine*, *26*(4), 217–238. https://doi.org/10.2165/00007256-199826040-00002

Barnes, K. R., & Kilding, A. E. (2015a). Running economy: Measurement, norms, and determining factors. *Sports Medicine - Open*, *1*(8), 1–15. https://doi.org/10.1186/s40798-015-0007-y

Barnes, K. R., & Kilding, A. E. (2015b). Strategies to Improve Running Economy. *Sports Medicine*, *45*(1), 37–56. https://doi.org/10.1007/s40279-014-0246-y

Bentley, D. J., McNaughton, L. R., & Batterham, A. M. (2001). Prolonged stage duration during incremental cycle exercise: Effects on the lactate threshold and onset of blood lactate accumulation. *European Journal of Applied Physiology*, *85*(3–4), 351–357. https://doi.org/10.1007/s004210100452

Bentley, D. J., Newell, J., & Bishop, D. (2007). Incremental Exercise Test Design and Analysis: Implications for Performance Diagnostics in Endurance Athletes. *Sports Medicine*, *37*(7), 575–586. https://doi.org/10.2165/00007256-200737070-00002

Blagrove, R. C., Howatson, G., & Hayes, P. R. (2017). Test–retest reliability of physiological parameters in elite junior distance runners following allometric scaling. *European Journal of Sport Science*, *17*(10), 1231–1240. https://doi.org/10.1080/17461391.2017.1364301

Bonaventura, J. M., Sharpe, K., Knight, E., Fuller, K. L., Tanner, R. K., & Gore, C. J. (2015). Reliability and accuracy of six hand-held blood lactate analysers. *Journal of Sports Science & Medicine*, *14*(1), 203–214.

Brady, A. J., & Egan, B. (2024). Acute Ingestion of a Ketone Monoester without Co-Ingestion of Carbohydrate Improves Running Economy in Male Endurance Runners.

Medicine & Science in Sports & Exercise, *56*(1), 134–142. https://doi.org/10.1249/MSS.0000000000003278

Brisswalter, J., & Legros, P. (1994). Daily stability in energy cost of running, respiratory parameters and stride rate among well-trained middle distance runners. *International Journal of Sports Medicine*, *15*(5), 238–241. https://doi.org/10.1055/s-2007-1021053

Burke, L. M. (2017). Practical Issues in Evidence-Based Use of Performance Supplements: Supplement Interactions, Repeated Use and Individual Responses. *Sports Medicine*, *47*(S1), 79–100. https://doi.org/10.1007/s40279-017-0687-1

Burke, L. M., Castell, L. M., Casa, D. J., Close, G. L., Costa, R. J. S., Desbrow, B., Halson, S. L., Lis, D. M., Melin, A. K., Peeling, P., Saunders, P. U., Slater, G. J., Sygo, J., Witard, O. C., Bermon, S., & Stellingwerff, T. (2019). International Association of Athletics Federations Consensus Statement 2019: Nutrition for Athletics. *International Journal of Sport Nutrition and Exercise Metabolism*, *29*(2), 73–84. https://doi.org/10.1123/ijsnem.2019-0065

Burnley, M., & Jones, A. M. (2007). Oxygen uptake kinetics as a determinant of sports performance. *European Journal of Sport Science*, *7*(2), 63–79. https://doi.org/10.1080/17461390701456148

Cicchetti, D.V., 1994. Guidelines, criteria, and rules of thumb for evaluating normed and standardized assessment instruments in psychology. Psychol. Assess. 6, 284–290. https://doi.org/10.1037/1040-3590.6.4.284

Currell, K., & Jeukendrup, A. E. (2008). Validity, Reliability and Sensitivity of Measures of Sporting Performance: *Sports Medicine*, *38*(4), 297–316. https://doi.org/10.2165/00007256-200838040-00003

Doma, K., Deakin, G. B., Leicht, A. S., & Sealey, R. M. (2012). The reliability of running economy among trained distance runners and field-based players. *Journal of Exercise Science and Fitness*, *10*(2), 90–96. https://doi.org/10.1016/j.jesf.2012.10.006

Foxdal, P., Sjödin, A., & Sjödin, B. (1996). Comparison of Blood Lactate Concentrations Obtained During Incremental and Constant Intensity Exercise. *International Journal of Sports Medicine*, *17*(05), 360–365. https://doi.org/10.1055/s-2007-972861

Gavin, J. P., Willems, M. E. T., & Myers, S. D. (2014). Reproducibility of lactate markers during 4 and 8min stage incremental running: A pilot study. *Journal of Science and Medicine in Sport*, *17*(6), 635–639. https://doi.org/10.1016/j.jsams.2013.08.006

Gullstrand, L., Sjüdin, B., & Svedenhag, J. (1994). Blood sampling during continuous running and 30-second intervals on a treadmill: Effects on the lactate threshold results? *Scandinavian Journal of Medicine & Science in Sports*, *4*(4), 239–242. https://doi.org/10.1111/j.1600-0838.1994.tb00434.x

Hopkins, W. G. (2000). Measures of Reliability in Sports Medicine and Science. *Sports Med*, *30*(1), 1–15. https://doi.org/10.2165/00007256-200030010-00001

Jamnick, N. A., Botella, J., Pyne, D. B., & Bishop, D. J. (2018). Manipulating graded exercise test variables affects the validity of the lactate threshold and V˙O2peak. *PLOS ONE*, *13*(7), e0199794. https://doi.org/10.1371/journal.pone.0199794

Jamnick, N. A., Pettitt, R. W., Granata, C., Pyne, D. B., & Bishop, D. J. (2020). An Examination and Critique of Current Methods to Determine Exercise Intensity. *Sports Medicine*, *50*(10), 1729–1756. https://doi.org/10.1007/s40279-020-01322-8

Jones, A. M. (2006). The Physiology of the World Record Holder for the Women's Marathon. *International Journal of Sports Science & Coaching*, *1*(2), 101–116. https://doi.org/10.1260/174795406777641258

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

Jones, A. M., Kirby, B. S., Clark, I. E., Rice, H. M., Fulkerson, E., Wylie, L. J., Wilkerson, D. P., Vanhatalo, A., & Wilkins, B. W. (2021). Physiological demands of running at 2-hour marathon race pace. *Journal of Applied Physiology*, *130*(2), 369–379. https://doi.org/10.1152/japplphysiol.00647.2020

Kuipers, H., Rietjens, G., Verstappen, F., Schoenmakers, H., & Hofman, G. (2003). Effects of Stage Duration in Incremental Running Tests on Physiological Variables. *International Journal of Sports Medicine*, *24*(7), 486–491. https://doi.org/10.1055/s-2003-42020

Lakens, D. (2013). Calculating and reporting effect sizes to facilitate cumulative science: A practical primer for t-tests and ANOVAs. *Frontiers in Psychology*, *4*(863), 1–12. https://doi.org/10.3389/fpsyg.2013.00863

Lambert, M. I., Mbambo, Z. H., & Gibson, A. S. C. (1998). Heart rate during training and competition for longdistance running. *Journal of Sports Sciences*, *16*, 85–90. https://doi.org/10.1080/026404198366713

Laursen, P. B., Francis, G. T., Abbiss, C. R., Newton, M. J., & Nosaka, K. (2007). Reliability of Time-to-Exhaustion versus Time-Trial Running Tests in Runners. *Medicine and Science in Sports and Exercise*, *39*(8), 1374–1379. https://doi.org/10.1249/mss.0b013e31806010f5

Lourenço, T. F., Martins, L. E. B., Tessutti, L. S., Brenzikofer, R., & Macedo, D. V. (2011). Reproducibility of an incremental treadmill VO(2)max test with gas exchange analysis for runners. *Journal of Strength and Conditioning Research*, *25*(7), 1994–1999. https://doi.org/10.1519/JSC.0b013e3181e501d6

McKay, A. K. A., Stellingwerff, T., Smith, E. S., Martin, D. T., Mujika, I., Goosey-Tolfrey, V. L., Sheppard, J., & Burke, L. M. (2022). Defining Training and Performance Caliber: A Participant Classification Framework. *International Journal of Sports Physiology and Performance*, *17*(2), 317–331. https://doi.org/10.1123/ijspp.2021-0451

Morgan, D. W., Martin, P. E., Krahenbuhl, G. S., & Baldini, F. D. (1991). Variability in running economy and mechanics among trained male runners. *Medicine and Science in Sports and Exercise*, *23*(3), 378–383. https://doi.org/10.1249/00005768-199103000-00018

Pereira, M. A., & Freedson, P. S. (1997). Intraindividual variation of running economy in highly trained and moderately trained males. *International Journal of Sports Medicine*, *18*(2), 118–124. https://doi.org/10.1055/s-2007-972606

Saunders, P. U., Pyne, D. B., Telford, R. D., & Hawley, J. A. (2004). Reliability and Variability of Running Economy in Elite Distance Runners: *Medicine & Science in Sports & Exercise*, *36*(11), 1972–1976. https://doi.org/10.1249/01.MSS.0000145468.17329.9F

Schaun, G. Z. (2017). The Maximal Oxygen Uptake Verification Phase: A Light at the End of the Tunnel? *Sports Medicine - Open*, *3*(1), 44. https://doi.org/10.1186/s40798-017-0112-1

Shaw, A. J., Ingham, S. A., Fudge, B. W., & Folland, J. P. (2013). The reliability of running economy expressed as oxygen cost and energy cost in trained distance runners. *Applied Physiology, Nutrition, and Metabolism*, *38*(12), 1268–1272. https://doi.org/10.1139/apnm-2013-0055

Sjödin, B., & Svedenhag, J. (1985). Applied Physiology of Marathon Running: *Sports Medicine*, *2*(2), 83–99. https://doi.org/10.2165/00007256-198502020-00002

Succi, P. J., Benitez, B., Kwak, M., & Bergstrom, H. C. (2023). Methodological considerations for the determination of VO2max in healthy men. *European Journal of Applied Physiology*, *123*(1), 191–199. https://doi.org/10.1007/s00421-022-05033-y

Thomas, D. T. T., Erdman, K. A., & Burke, L. M. (2016). Nutrition and Athletic Performance: *Medicine & Science in Sports & Exercise*, *48*(3), 543–568. https://doi.org/10.1249/MSS.0000000000000852

Williams, T. J., Krahenbuhl, G. S., & Morgan, D. W. (1991). Daily variation in running economy of moderately trained male runners. *Medicine and Science in Sports and Exercise*, *23*(8), 944–948. https://doi.org/10.1249/00005768-199108000-00010

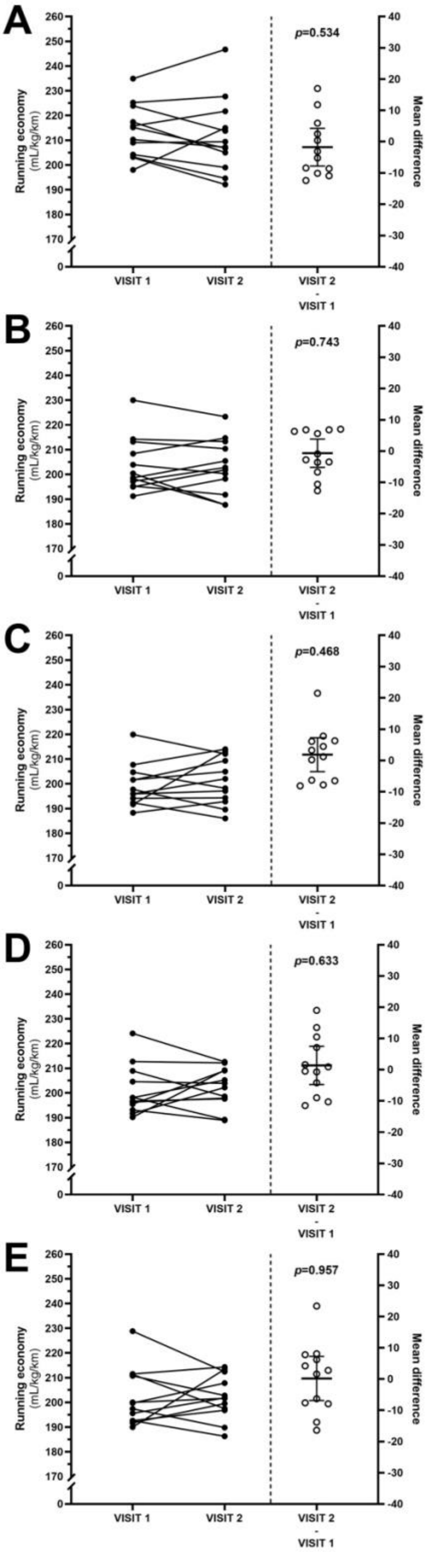

Figure 1. Pairwise comparison of running economy (RE) between visit 1 and visit 2 and the mean difference at 10 km·h^{-1} (A), 11 km·h^{-1} (B), 12 km·h^{-1} (C), 13 km·h^{-1} (D), and 14 km·h^{-1} (E). Filled circles represent the individual data points from each visit (left); Clear circles represent the individual difference between visits with data presented as mean and error bars representing 95% CI (right).

Table 1. Descriptive analysis and test-retest reliability of RE, RPE, HR, $\dot{V}O_{2peak}$, and TTE at each stage of a two-phase incremental treadmill test.

	Stage	Visit 1	Visit 2	CV (%)	95% CI of CV (%)	ICC	95% CI of ICC	SWC[a]	P	ES
RE ($mL \cdot kg \cdot km^{-1}$)	10 $km \cdot h^{-1}$	213.3±10.9	211.6±15.2	2.7	1.8 – 3.6	0.74	0.32 – 0.92	2.6	0.534	0.12
	11 $km \cdot h^{-1}$	203.8±11.0	203.1±11.0	2.2	1.5 – 2.8	0.79	0.42 – 0.94	2.2	0.743	0.06
	12 $km \cdot h^{-1}$	199.3±8.6	201.1±9.6	2.3	1.3 – 3.4	0.56	0.02 – 0.85	1.8	0.468	0.19
	13 $km \cdot h^{-1}$	200.8±10.0	202.2±8.1	2.6	1.4 – 3.8	0.44	-0.16 – 0.80	1.8	0.633	0.15
	14 $km \cdot h^{-1}$	201.8±11.6	202.0±8.9	3.1	1.9 – 4.3	0.42	-0.17 – 0.79	2.1	0.957	0.02
RPE (AU)	10 $km \cdot h^{-1}$	8.8±1.4	7.7±1.6	10.8	5.2 – 16.4	0.60	0.07 – 0.86	0.3	0.015	0.71
	11 $km \cdot h^{-1}$	10.3±1.3	9.6±1.4	6.1	2.4 – 9.7	0.68	0.20 – 0.90	0.3	0.054	0.50
	12 $km \cdot h^{-1}$	11.8±1.1	11.3±1.5	6.5	3.2 – 9.7	0.58	0.04 – 0.86	0.3	0.191	0.37
	13 $km \cdot h^{-1}$	13.3±1.4	12.7±1.6	4.3	1.1 – 7.5	0.74	0.31 – 0.92	0.3	0.089	0.39
	14 $km \cdot h^{-1}$	14.8±1.8	14.6±1.6	2.3	0.6 – 4.0	0.88	0.62 – 0.96	0.3	0.504	0.11
HR ($beats \cdot min^{-1}$)	10 $km \cdot h^{-1}$	127±10	122±10	3.3	1.9 – 4.6	0.89	0.67 – 0.97	2.0	0.002	0.48
	11 $km \cdot h^{-1}$	137±10	131±11	3.2	2.1 – 4.4	0.94	0.80 – 0.98	2.1	<0.001	0.55
	12 $km \cdot h^{-1}$	145±11	140±12	2.7	1.7 – 3.7	0.95	0.82 – 0.98	2.3	0.001	0.42
	13 $km \cdot h^{-1}$	156±12	151±11	2.0	1.0 – 3.1	0.93	0.77 – 0.98	2.3	0.006	0.42
	14 $km \cdot h^{-1}$	164±12	160±11	2.0	1.0 – 3.0	0.93	0.76 – 0.98	2.4	0.006	0.34
	Peak	182±12	181±13	1.1	0.7 – 1.5	0.98	0.91 – 0.99	2.5	0.058	0.08
$\dot{V}O_2$ ($mL \cdot kg \cdot min^{-1}$)	Peak	59.6±4.9	60.0±5.7	2.6	1.2 – 4.0	0.82	0.46 – 0.95	1.1	0.727	0.07
TTE (s)	Peak	375±79	377±83	7.2	2.8 – 11.6	0.76	0.33 – 0.93	16.3	0.508	0.02

Data are presented as mean±SD. CV, intraindividual variation expressed as a coefficient of variation; 95% CI, 95% confidence interval for the coefficient of variation or intraclass correlation coefficient; ICC, intraclass correlation coefficient; SWC, smallest worthwhile change; ES, effect size calculated as Hedges' *g*; $\dot{V}O_2$, oxygen consumption; RE, running economy; RPE, ratings of perceived exertion; HR, heart rate; TTE, time to exhaustion. [a]SWC is presented in the units of measure indicated in the leftmost column for each parameter.

Table 2. Descriptive analysis and test-retest reliability of $\dot{V}E$, $\dot{V}O_2$, $\dot{V}CO_2$, and RER at each stage of a two-phase incremental treadmill test.

	Stage	Visit 1	Visit 2	CV (%)	95% CI of CV (%)	ICC	95% CI of ICC	SWC[a]	P	ES
$\dot{V}E$ ($L \cdot min^{-1}$)	10 $km \cdot h^{-1}$	67.5±7.3	66.0±8.6	3.5	2.3 – 4.8	0.89	0.68 – 0.97	1.6	0.172	0.18
	11 $km \cdot h^{-1}$	73.1±7.7	72.4±8.7	2.6	1.4 – 3.7	0.91	0.71 – 0.97	1.6	0.454	0.08
	12 $km \cdot h^{-1}$	78.8±9.7	78.8±9.4	2.1	1.5 – 2.7	0.96	0.86 – 0.99	1.9	0.968	0.00
	13 $km \cdot h^{-1}$	88.0±10.5	88.3±11.2	2.2	1.4 – 3.0	0.95	0.84 – 0.99	2.2	0.802	0.03
	14 $km \cdot h^{-1}$	98.9±13.6	96.6±13.7	3.6	2.5 – 4.8	0.92	0.75 – 0.98	2.7	0.163	0.16
	Peak	153.0±15.2	154.5±15.8	1.8	0.8 – 2.7	0.95	0.82 – 0.99	3.1	0.340	0.09
$\dot{V}O_2$ ($L \cdot min^{-1}$)	10 $km \cdot h^{-1}$	2.670±0.283	2.641±0.335	2.5	1.7 – 3.4	0.94	0.80 – 0.98	0.062	0.371	0.09
	11 $km \cdot h^{-1}$	2.807±0.309	2.786±0.297	2.2	1.5 – 3.0	0.94	0.81 – 0.98	0.061	0.509	0.07
	12 $km \cdot h^{-1}$	2.994±0.311	3.009±0.292	2.4	1.4 – 3.4	0.90	0.70 – 0.97	0.060	0.709	0.05
	13 $km \cdot h^{-1}$	3.273±0.381	3.279±0.333	2.7	1.5 – 3.8	0.90	0.68 – 0.97	0.071	0.899	0.02
	14 $km \cdot h^{-1}$	3.539±0.409	3.528±0.368	3.1	1.8 – 4.3	0.86	0.59 – 0.96	0.078	0.860	0.03
	Peak	4.465±0.499	4.493±0.522	2.3	1.2 – 3.4	0.92	0.72 – 0.98	0.102	0.666	0.05
$\dot{V}CO_2$ ($L \cdot min^{-1}$)	10 $km \cdot h^{-1}$	2.633±0.324	2.631±0.356	2.4	1.4 – 3.4	0.94	0.82 – 0.98	0.068	0.950	0.01
	11 $km \cdot h^{-1}$	2.782±0.341	2.814±0.297	3.0	1.9 – 4.1	0.89	0.67 – 0.97	0.064	0.466	0.10
	12 $km \cdot h^{-1}$	2.959±0.338	3.030±0.303	4.2	2.9 – 5.5	0.80	0.44 – 0.94	0.064	0.253	0.21
	13 $km \cdot h^{-1}$	3.239±0.405	3.303±0.345	3.7	2.3 – 5.1	0.85	0.56 – 0.96	0.075	0.303	0.16
	14 $km \cdot h^{-1}$	3.545±0.468	3.548±0.374	4.3	2.5 – 6.0	0.79	0.41 – 0.93	0.085	0.968	0.01
	Peak	5.176±0.536	5.294±0.515	2.8	1.3 – 4.3	0.87	0.58 – 0.96	0.105	0.181	0.22
RER ($\dot{V}CO_2/\dot{V}O_2$)	10 $km \cdot h^{-1}$	0.98±0.04	1.00±0.03	1.3	0.5 – 2.0	0.80	0.44 – 0.94	0.01	0.127	0.55
	11 $km \cdot h^{-1}$	0.99±0.04	1.01±0.03	1.9	1.2 – 2.6	0.62	0.10 – 0.87	0.01	0.067	0.55
	12 $km \cdot h^{-1}$	0.99±0.04	1.01±0.04	2.7	1.9 – 3.5	0.50	-0.07 – 0.83	0.01	0.144	0.48
	13 $km \cdot h^{-1}$	0.99±0.04	1.01±0.05	1.8	0.7 – 2.8	0.74	0.31 – 0.92	0.01	0.143	0.43
	14 $km \cdot h^{-1}$	1.00±0.05	1.01±0.05	2.0	1.2 – 2.8	0.72	0.27 – 0.91	0.01	0.754	0.19
	Peak	1.19±0.05	1.20±0.06	1.3	0.7 – 1.9	0.89	0.64 – 0.97	0.01	0.142	0.17

Data are presented as mean±SD. CV, intraindividual variation expressed as a coefficient of variation; 95% CI, 95% confidence interval for the coefficient of variation or intraclass correlation coefficient; ICC, intraclass correlation coefficient; SWC, smallest worthwhile change; ES, effect size calculated as Hedges' *g*; $\dot{V}E$, minute ventilation; $\dot{V}O_2$, oxygen consumption; $\dot{V}CO_2$, carbon dioxide production; RER, respiratory exchange ratio. [a]SWC is presented in the units of measure indicated in the leftmost column for each parameter.

Table 3. Descriptive analysis and test-retest reliability of blood glucose and lactate at each stage of a two-phase incremental treadmill test.

	Stage	Visit 1	Visit 2	CV (%)	95% CI of CV (%)	ICC	95% CI of ICC	SWC[a]	P	ES
Blood glucose (mM)	-20 min	5.0±0.3	5.2±0.3	4.0	2.2 – 5.8	0.29	-0.31 – 0.73	0.1	0.121	0.64
	0 min	8.9±1.4	9.2±1.8	9.5	5.9 – 13.1	0.59	0.05 – 0.86	0.3	0.384	0.18
	10 $km \cdot h^{-1}$	5.8±1.0	6.1±1.5	11.7	6.4 – 16.9	0.48	-0.10 – 0.82	0.3	0.580	0.23
	11 $km \cdot h^{-1}$	3.9±0.7	4.0±1.3	15.2	9.2 – 21.1	0.38	-0.22 – 0.77	0.2	0.822	0.09
	12 $km \cdot h^{-1}$	4.1±0.5	4.2±0.8	12.2	7.6 – 16.9	0.23	-0.37 – 0.70	0.1	0.612	0.14
	13 $km \cdot h^{-1}$	5.0±0.8	5.0±0.8	10.5	5.7 – 15.4	0.33	-0.28 – 0.74	0.2	0.952	0.00
	14 $km \cdot h^{-1}$	5.8±0.9	5.7±0.9	9.1	5.7 – 12.5	0.51	-0.05 – 0.83	0.2	0.727	0.11
	Peak	7.0±0.7	6.8±1.1	10.1	6.0 – 14.2	0.11	-0.50 – 0.65	0.2	0.420	0.21
Blood lactate (mM)	-20 min	0.9±0.1	1.1±0.2	14.6	8.6 – 20.6	0.23	-0.37 – 0.70	<0.1	0.162	1.22
	0 min	1.1±0.3	1.2±0.3	17.8	12.7 – 23.0	0.58	0.04 – 0.86	0.1	0.317	0.32
	10 $km \cdot h^{-1}$	1.2±0.4	1.4±0.5	16.9	9.0 – 24.8	0.71	0.27 – 0.91	0.1	0.051	0.43
	11 $km \cdot h^{-1}$	1.1±0.4	1.3±0.4	14.0	6.3 – 21.6	0.82	0.49 – 0.95	0.1	0.036	0.48
	12 $km \cdot h^{-1}$	1.2±0.5	1.4±0.5	12.1	6.9 – 17.2	0.93	0.78 – 0.98	0.1	0.009	0.39
	13 $km \cdot h^{-1}$	1.6±0.8	1.7±0.7	17.1	7.7 – 26.5	0.80	0.45 – 0.94	0.1	0.759	0.13
	14 $km \cdot h^{-1}$	2.2±1.4	2.2±1.3	11.7	6.0 – 17.4	0.95	0.84 – 0.99	0.3	1.000	0.00
	Peak	7.5±1.6	6.9±1.7	11.8	6.2 – 17.4	0.62	0.07 – 0.88	0.3	0.198	0.35

Data are presented as mean±SD. CV, intraindividual variation expressed as a coefficient of variation; 95% CI, 95% confidence interval for the coefficient of variation or intraclass correlation coefficient; ICC, intraclass correlation coefficient; SWC, smallest worthwhile change; ES, effect size calculated as Hedges' *g*. [a]SWC is presented in the units of measure indicated in the leftmost column for each parameter.