



ELSEVIER

---

**Journal of  
Science and  
Medicine in  
Sport**

---

www.elsevier.com/locate/jsams

ORIGINAL PAPER

# The reliability of physiological and performance measures during simulated team-sport running on a non-motorised treadmill

Anita C. Sirotic\*, Aaron J. Coutts

*School of Leisure, Sport & Tourism, University of Technology, Sydney, Australia*

Received 14 August 2006; received in revised form 11 April 2007; accepted 11 April 2007

**KEYWORDS**Reproducibility;  
Ratio limits of  
agreement;  
Team-sport;  
Match simulation

**Summary** The aim of this study was to determine the reliability of a non-motorised treadmill team-sport simulation for measuring physiological responses and performance demands of team sports. Following familiarisation, 11 team-sport athletes completed a peak sprinting speed assessment followed by a 30-min team-sport simulation on the non-motorised treadmill, on three occasions, 5 days apart. Several performance (total distance, distance covered during each speed category, total work, high-intensity activity, mean maximal sprinting speed and power) and physiological variables ( $\dot{V}_{O_2}$ , heart rate and blood measures) were measured. A one-way analysis of variance and ratio limits of agreement were used to compare the results from each trial. Significant differences were established in total sprint distance and high-intensity activity between trials 1–2 and trials 1–3 and 3-s mean maximal sprinting speed for trials 1–3 ( $p < 0.05$ ). No other significant differences were identified. Moderate to high intraclass correlation coefficients (i.e.,  $>0.8$ ) were identified in 11 of the 18 physiological and performance variables measured. Ratio limits of agreement for total distance covered and total work performed during the team-sport simulation were 0.99 ( $*/\div 1.05$ ) and 0.97 ( $*/\div 1.09$ ), respectively. Largest measurement error was shown in post-exercise blood lactate concentration with a coefficient of variation of 17.6%. All other measures showed low coefficients of variation of  $\leq 10\%$ . These results show that the non-motorised treadmill team-sport simulation provides a reliable tool for assessing and monitoring physiological and performance demands of team-sport activity. We recommend the inclusion of two familiarisation sessions prior to testing.

© 2007 Sports Medicine Australia. Published by Elsevier Ltd. All rights reserved.

## Introduction

Team sports with similar match running demands<sup>1</sup> require the combination of repeated high-intensity bouts of exercise interspersed with longer periods

\* Corresponding author.

E-mail address: Anita.C.Sirotic@uts.edu.au (A.C. Sirotic).

of low-intensity exercise for a duration of between 30 min and 120 min, depending on player position and interchange.<sup>2</sup> In response, several team-sport simulations using a non-motorised treadmill (NMT) have been developed.<sup>3–5</sup> It has been suggested that these team-sport simulations can be consistently reproduced on a NMT in a controlled environment, thereby providing a useful tool for monitoring running performance and physiological changes specific to team sports. To date, the 'technological and biological' error of the performance outcomes and physiological responses of these team-sport simulations have not been extensively determined.

The NMT allows a close replication of the physiological workload and running demands of a team-sport match as it enables near maximal velocities to be obtained, allows for instantaneous changes in running speed and provides real-time measures of power output.<sup>6</sup> These characteristics are vital for accurate monitoring of the physical demands of team-sport activity in a laboratory. Additionally, the controlled laboratory environment also allows for physiological (i.e.,  $\dot{V}_{O_2}$  and blood measures) and performance variables (i.e., power output) to be continually measured. At present these variables are difficult to measure during competitive matches as the collecting procedures interfere with normal play.<sup>7</sup> Other advantages of using a controlled laboratory setting is that the influence of the opposition, environment, team-tactics, match score or officials is removed. Collectively, these standardised conditions should allow an increased level of reliability of the results taken from this test.

To date, several studies have determined the reproducibility of brief repeated sprint efforts on the NMT.<sup>8,9</sup> However, while NMT team-sport simulations have been validated,<sup>3,4</sup> the comprehensive reliability of such simulations is still unknown. To our knowledge, there are only two studies that have determined the reliability of some physiological and performance measures of a NMT team-sport simulation. The first study determined the reproducibility of mean HR during the first and second half of a 90-min NMT soccer-specific protocol.<sup>3</sup> The second study determined the reliability of the total distance (TD) covered during a 15-min period of a 90-min soccer simulation on a NMT.<sup>10</sup> In these studies, no other performance or physiological measures were examined for reproducibility.

Reliability of performance measures and physiological responses to NMT team-sport simulations needs to be assessed. The information obtained from these reliability measures will allow sports sci-

entists to be able to interpret 'real' changes, or changes independent of 'technological and biological' error.<sup>11</sup> The ability to identify a 'real' change in these measures will enable sports scientists to assess more accurately the effects of intervention such as training strategies on team-sport running performance and physiological responses. Therefore, the purpose of this study is to report on the reliability of running performance and physiological responses to a generic team-sport protocol on a NMT.

## Methods

### Subjects

Eleven moderately trained (age =  $23.6 \pm 4.5$  years; and  $\dot{V}_{O_2 \max} = 52.7 \pm 4.5$  mL kg<sup>-1</sup> min<sup>-1</sup>) male team-sport athletes participated in this study. All subjects were requested to complete their normal training programs during the testing period. Subjects were also instructed to abstain from physical training and products containing caffeine in the 24 h prior and from consuming food in the 2 h before each test. During the 24–48 h prior to each testing session subjects were asked to complete only low-volume, low-intensity training. Additionally, subjects were asked to standardise food and fluid intake and training practices during the 48 h prior to each test. Subjects recorded food and fluid intake and training practices so that replication of the same regimen could be practised for all testing occasions. Prior to commencing testing, written informed consent was voluntarily obtained by all subjects. Ethical approval was granted by the University Human Research Ethics Committee for all experimental procedures.

### Training history

All subjects were regularly participating in team sports (four soccer, four Australian Football and three rugby union) for at least 12 months prior to the study. Subjects were in the competitive phase of the training season while undergoing all testing procedures. During this competitive phase of training, each subject completed between 1 and 3 team training sessions as well as 3 and 5 individual physical training sessions per week consisting of interval training, distance running, sprint training and/or resistance training. Subjects also participated in one competitive match per week in their respective team sport.

## Experimental design

Subjects visited the laboratory on five occasions, each separated by 5 days. Visit 1 included familiarisation with testing procedures. Subjects completed a  $\dot{V}_{O_2 \max}$  test on visit 2. A peak sprinting speed assessment (PSSA) and 30-min NMT team-sport simulation were performed on visits 3–5. On each occasion the tests were performed at the same time of day (within 2 h) to minimise diurnal variations.

## Maximal oxygen uptake

Maximal oxygen uptake ( $\dot{V}_{O_2 \max}$ ) was determined using a discontinuous, incremental treadmill run to exhaustion using open-circuit spirometry (Physio-Dyne® Gas Analysis System, Quogue, NY, USA). The work protocol and criteria for attainment of  $\dot{V}_{O_2 \max}$  used in this study have been previously described.<sup>12</sup>

## Treadmill design

The treadmill design was a modified version of the original NMT system designed by Lakomy.<sup>6</sup> The NMT (Force Tread Dynameter, Woodway, USA) consisted of an endless wire reinforced lateral belt that was attached to a metal frame that included a handrail, display unit and vertical metal strut with sliding gauge. A 100-kg horizontal 'S' beam load cell (Model number: HTC-500SS, Liftcells, Adelaide, Australia) was attached to the vertical strut via the sliding gauge, which locked in place to ensure no movement during testing. The sliding gauge allowed the horizontal load cell to be adjusted vertically. A 'Y' jointed steel wire attached the horizontal load cell to a tether belt worn around the waist of each subject. To minimise errors in force measurement, the tether belt was adjusted to 8° above the horizontal of the subject's waist prior to testing using an electronic level (Smart Tool, Macklangburg-Duncan, Oklahoma, USA).<sup>6</sup> The load cell was calibrated before and after each test using a range of known weights. Treadmill belt speed was monitored by two optical speed photomicrosensors (Model number: EE-SX670, Omron Electronics, Schaumburg, IL) mounted on the rear roller shaft of the treadmill belt.

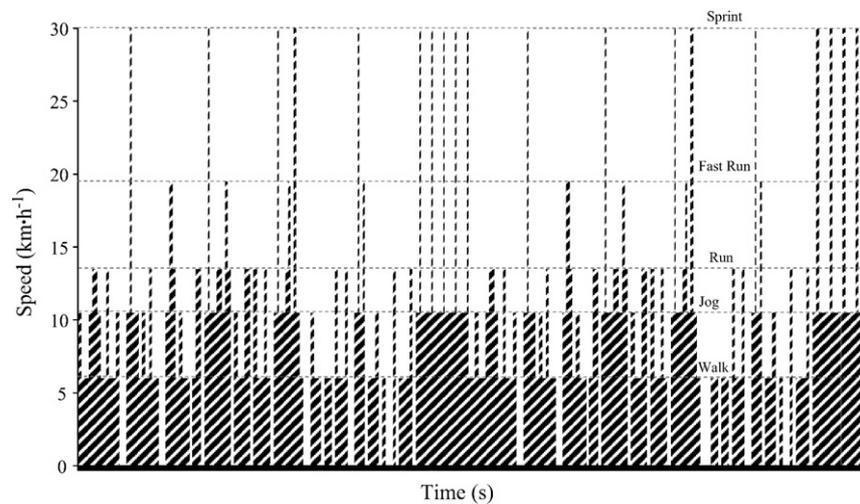
Treadmill belt speed, distance and horizontal forces were collected at a sampling rate of 10 Hz via the XPV7 PCB interface (Fitness Technology, Adelaide, Australia) and analysed using the Force 3.0 Software (Innervations Software, Joondalup, Australia). Data collected were exported to Microsoft Excel® (Microsoft, Redmond, USA) where the data

were synchronised with the team-sport simulation starting time. The horizontal power output was calculated from the product of treadmill belt speed ( $\text{km h}^{-1}$ ) and horizontal force (N) measured from the horizontal load cell.

## Team-sport simulation

The development of the team-sport simulation has previously been described.<sup>4</sup> The simulation was designed to mimic the running work profile of several team sports and was based on time–motion data of sports such as soccer, rugby league, rugby union and Australian Football.<sup>13–16</sup> Two 15-min activity profiles were performed succinctly throughout the team-sport simulation for a total duration of 30 min. Included in these activity profiles were six speed categories: standing (0% of maximal sprinting speed (MSS)); walking (20% of MSS); jogging (35% of MSS); running (45% of MSS); fast running (65% of MSS); and sprinting (100% of MSS). These six speed categories were divided into aerobic (walking, jogging and running) and anaerobic (fast running and sprinting) activities based on the intensity of the movement. The six speed categories were also designated a particular duration based on time–motion data from team sports.<sup>13</sup> The time duration assigned and total time spent in each speed category during the team-sport simulation are similar to those previously described.<sup>4</sup> A total of 181 activity changes were included in the team-sport simulation, with a change of activity occurring on average once every 10 s. This frequency of activity changes is slightly lower when compared to time–motion data of a variety of team sports.<sup>17–19</sup> Fig. 1 shows the activity profile for a subject with a MSS of  $30 \text{ km h}^{-1}$ . During the team-sport simulation all effort was made to provide equal verbal encouragement to each subject to sprint as hard as possible when required. At all other times each subject was verbally instructed to match his speed with the target speed displayed.

Prior to the commencement of the team-sport simulation, MSS was determined during the PSSA. After a 5-min standardised warm-up on the NMT, MSS was measured using previously reported methods.<sup>4,5</sup> Each subject completed three maximal 3-s and 6-s sprints, alternately, separated by 2 min of active recovery. MSS was calculated as the highest speed obtained in a single second from the 3-s and 6-s sprints. Mean MSS (mMSS), mean maximal power (mMP) and distance covered during the best 3-s and 6-s sprint from trial 1, trial 2 and trial 3 were determined to assess the reliability of NMT sprinting.



**Figure 1** The 30-min team-sport running simulation activity profile for a subject with a maximal speed of  $30 \text{ km h}^{-1}$ . Two 15 min periods were completed with a 2-min recovery after completion of the first 15-min activity profile.

### Physiological responses

Continuous oxygen uptake was measured during the team-sport simulation using open-circuit spirometry, which was calibrated before and after each test with reference and calibration gases of known concentrations. The pneumotach was calibrated with ambient air using a 3-L syringe (Hans Rudolph, Inc., Kansas City, USA). Heart rate (HR) was recorded every 5 s during the team-sport simulation using Polar Team System HR monitors (Polar, OY, Finland) and analysed with the Polar Precision Performance SW Version 4.02 software package (Polar, OY, Finland). Blood lactate concentration was determined from 25- $\mu\text{L}$  capillarised blood samples taken from hyperaemic fingertips. Samples were drawn from all subjects immediately following the team-sport simulation and analysed using an Accusport Portable Lactate Analyser (Boehringer Mannheim, Germany) immediately following collection.

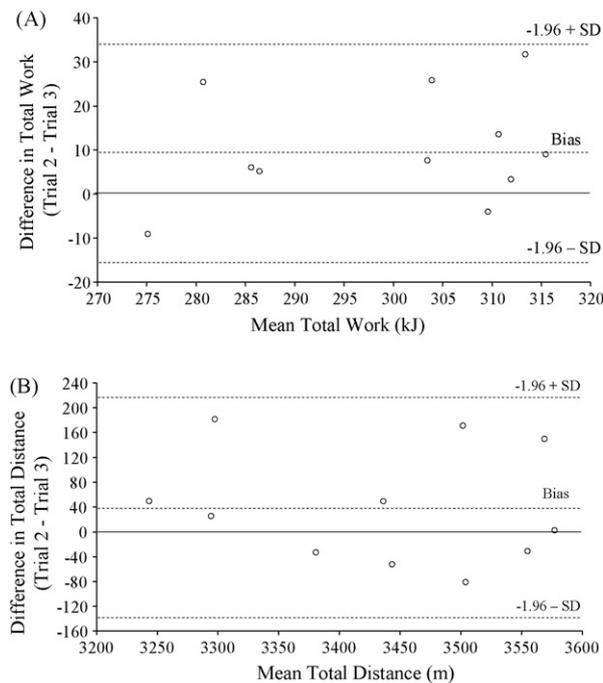
### Statistical analyses

In order to allow comparison between previous work and the results of the present study, a variety of reliability measures were used. Test–retest differences were tested for normality using the Shapiro–Wilks statistic. If test–retest differences were not normally distributed, logarithmic transformation was performed and these data sets were retested. A one-way analysis of variance (ANOVA) with repeated measures was used to detect any differences in measurement variables between the three trials. Where appropriate, post hoc comparisons (least significance difference) were employed. Mauchly’s test of sphericity was used

to assess homogeneity of variance. A Greenhouse–Geisser adjustment was used if assumptions of homogeneity were violated. As heteroscedasticity was present in some data, logarithmic transformation was performed. The coefficient of variation (CV), intraclass correlation coefficient (ICC), 95% confidence intervals (CI), typical error (TE) and technical error or measurement (TEM) were calculated according to Hopkins.<sup>20</sup> Ratio measures for 95% limits of agreement were calculated as recommended by Bland and Altman.<sup>21</sup> Analyses were performed using Microsoft Excel® (Microsoft, Redmond, USA) and SPSS (Version 14.0, Chicago, USA). Statistical significance was set at  $p < 0.05$ . All data are reported as the mean  $\pm$  SD unless otherwise stated.

### Results

Significant differences were identified in three of the 18 variables measured during the team-sport simulation. Total sprint distance (SD) and high-intensity activity (HIA) were significantly different in trials 1–2 ( $\text{SD} = 407.12 \pm 26.37 \text{ m}$  versus  $417.98 \pm 18.16 \text{ m}$ ,  $\text{HIA} = 571.36 \pm 35.66 \text{ m}$  versus  $585.93 \pm 23.12 \text{ m}$ ,  $p < 0.05$ ) and trials 1–3 ( $\text{SD} = 407.12 \pm 26.37 \text{ m}$  versus  $424.73 \pm 18.80 \text{ m}$ ,  $\text{HIA} = 571.36 \pm 35.66 \text{ m}$  versus  $592.34 \pm 24.64 \text{ m}$ ,  $p < 0.05$ ). A significant difference was also shown in 3-s mSS between trials 1 and 3 ( $7.05 \pm 0.41 \text{ m s}^{-1}$  versus  $7.24 \pm 0.44 \text{ m s}^{-1}$ ,  $p < 0.05$ ). Heteroscedasticity was present in total work (TW) performed and  $\dot{V}\text{O}_2$  during the team-sport simulation, 3-s SD and mMP and 6-s mSS data. Since heteroscedasticity indicates a linear relationship between the amount



**Figure 2** Bland–Altman plot of the total work performed (A) and total distance covered (B) during the NMT team-sport simulation from trial 2 and trial 3 ( $n = 11$ ).

of random error and the size of the measured value (i.e., that a higher absolute measurement error exists in subjects scoring highest values), all results were log transformed and expressed as ratio limits of agreement.<sup>11</sup>

### Reliability of performance measures

Table 1 shows the grand mean, CV, ratio limits of agreement, ICC and respective CI for performance variables measured during the team-sport simulation. Fig. 2 shows the Bland–Altman plot of the mean TW and TD for trials 2–3, respectively. Ratio limits of agreement for the TD covered during the team-sport simulation were similar for trials 1–2 and trials 2–3, respectively ( $1.01^*/\div 1.06$  versus  $0.99^*/\div 1.05$ ). Ratio limits of agreement of mMP for trials 2–3 were lower for a 6-s sprint ( $1.00^*/\div 1.16$ ) compared to a 3-s sprint ( $1.04^*/\div 1.30$ ). Based on the nomogram designed by Atkinson et al.<sup>22</sup> and using the ratio limits of agreement from trials 2 to 3, it was estimated that a sample size of between 5 and 9 is required to detect a 10% change in all performance measures described in Tables 1 and 2 (statistical power = 0.9). However, a sample size of 25 is needed to detect a 10% change in mMP during a 3-s sprint.

### Reliability of physiological measures

Table 3 shows the grand mean, CV, ratio limits of agreement, ICC and respective CI for physiological variables measured during the team-sport simulation. Largest measurement error was shown in post-exercise  $[BLa^-]$  for both trials 1–2 and trials 2–3, respectively ( $1.12^*/\div 1.83$  versus  $0.97^*/\div 1.57$ ). To detect a 10% change in mean HR and  $\dot{V}O_2$  during the team-sport simulation and post-exercise  $[BLa^-]$ , a sample size of 5, 9 and 86 is required, respectively.<sup>22</sup>

### Discussion

From consideration of all reliability analyses performed, the results demonstrate that the NMT system and team-sport simulation provide a reliable tool for measuring performance demands and most physiological responses of various team sports in moderately trained athletes. The reliability established across the three trials revealed a high reproducibility, with the majority of variables reporting a CV of <5%. Out of the four remaining variables having a CV > 5%, three of these measures remained at a CV of  $\leq 10\%$ . It has been suggested that a CV of 10% is the criterion value commonly used to define an acceptable level of reliability in a test.<sup>22</sup> Another common criterion used to verify a reliable test is an ICC > 0.8.<sup>22</sup> Moderate to high ICC (i.e., >0.8) were identified in 11 of the 18 physiological and performance variables measured during the team-sport simulation, providing support to the reliability of the NMT for assessing team-sport running performance.<sup>23</sup>

In agreement with previous research,<sup>8,9</sup> the ratio limits of agreement for mSS in a 3-s and 6-s sprint in the present study are low for both bias ( $\leq 1.03$ ) and random error ( $\leq 1.05$ ). The low bias ( $\leq 1.04$ ) reported for mMP for the 3-s and 6-s sprint and random error (1.16) for the 6-s sprint in the present study are also similar to previous results of mMP for a 6-s sprint.<sup>8</sup> However, the random error in mMP for the 3-s sprint in the present study is slightly greater than the mMP for a 6-s sprint previously reported (1.30 versus 1.21).<sup>8</sup> The present results of a larger CV for mMP (10.1% versus 5.4%) and mSS (1.7% versus 1.3%) for a 3-s sprint compared to a 6-s sprint, respectively, show that, while a 3-s sprint on the NMT is reliable, a 6-s sprint on the NMT has even greater reliability. Therefore, on the basis of these results, we suggest that a 6-s sprint should be used to assess sprint performance on the NMT, especially when small changes in sprinting performance (i.e.,  $\geq 1\%$ ) are required.

**Table 1** Measures of reliability (with 95% confidence intervals) for total work and distances covered during each speed category from three trials of a 30-min NMT team-sport running simulation ( $n = 11$ )

Performance variable	Grand mean	CV (%)	CI (%)	Ratio limits of agreement	Lower–upper limit	ICC	CI
Trials 1–2							
Total work (kJ)	299.66	4.51	(3.20–8.24)	1.035*/÷ 1.130	247.35–350.35	0.203	(–0.452 to 0.716)
Total distance (m)	3437.62	2.21	(1.56–3.96)	1.011*/÷ 1.063	3270.28–3692.36	0.617	(0.028 to 0.888)
High-intensity activity (m)	578.75	2.34	(1.65–4.19)	1.026*/÷ 1.066	557.10–633.28	0.837	(0.477 to 0.957)
Walk (m)	706.28	3.29	(2.33–5.95)	1.019*/÷ 1.094	658.15–787.67	0.195	(–0.495 to 0.711)
Jog (m)	1390.75	2.12	(1.49–3.78)	1.017*/÷ 1.060	1334.81–1499.06	0.788	(0.356 to 0.942)
Run (m)	645.45	2.15	(1.51–3.84)	1.010*/÷ 1.061	614.39–696.18	0.827	(0.450 to 0.954)
Fast run (m)	166.19	2.36	(1.66–4.23)	1.022*/÷ 1.067	159.23–181.20	0.876	(0.583 to 0.968)
Sprint (m)	412.56	2.48	(1.75–4.45)	1.028*/÷ 1.070	396.22–453.83	0.835	(0.471 to 0.956)
Trials 2–3							
Total work (kJ)	299.66	3.06	(2.16–5.52)	0.966*/÷ 1.087	266.29–314.74	0.686	(0.146 to 0.911)
Total distance (m)	3436.48	1.91	(1.34–3.41)	0.988*/÷ 1.054	3223.48–3579.73	0.738	(0.247 to 0.927)
High-intensity activity (m)	589.14	1.46	(1.02–2.59)	1.011*/÷ 1.041	572.11–619.92	0.874	(0.577 to 0.967)
Walk (m)	707.15	3.63	(2.57–6.57)	0.983*/÷ 1.104	629.86–767.40	0.182	(–0.469 to 0.705)
Jog (m)	1392.51	2.84	(2.00–5.10)	0.985*/÷ 1.081	1269.64–1482.59	0.680	(0.136 to 0.909)
Run (m)	644.82	1.68	(1.18–3.00)	0.989*/÷ 1.047	608.94–668.02	0.854	(0.522 to 0.961)
Fast run (m)	167.78	1.73	(1.21–3.07)	0.998*/÷ 1.049	159.73–175.62	0.907	(0.672 to 0.976)
Sprint (m)	421.36	1.83	(1.29–3.27)	1.016*/÷ 1.052	407.12–450.23	0.836	(0.472 to 0.956)

CV: coefficient of variation; CI: confidence intervals; and ICC: intraclass correlation coefficient.

**Table 2** Measures of reliability (with 95% confidence intervals) for 3-s and 6-s sprint performance measures from three trials of a peak sprinting speed assessment on a NMT ( $n = 11$ )

Sprint variable	Grand mean	CV (%)	CI (%)	Ratio limits of agreement	Lower–upper limit	ICC	CI
<b>Trials 1–2</b>							
6-s distance (m)	37.23	2.45	(1.73–4.40)	1.018 * / ÷ 1.070	35.45–40.55	0.851	(0.512 to 0.960)
6-s mMP (W)	964.36	9.01	(6.50–17.14)	1.002 * / ÷ 1.270	760.72–1227.37	0.811	(0.410 to 0.949)
6-s mMSS ( $\text{m s}^{-1}$ )	7.29	1.34	(0.94–2.38)	1.005 * / ÷ 1.038	7.06–7.60	0.941	(0.785 to 0.985)
3-s distance (m)	17.87	6.16	(4.40–11.42)	1.048 * / ÷ 1.181	15.87–22.10	0.576	(–0.036 to 0.874)
3-s mMP (W)	1029.35	15.59	(11.51–31.46)	0.993 * / ÷ 1.494	683.88–1526.57	0.493	(–0.151 to 0.8.44)
3-s mMSS ( $\text{m s}^{-1}$ )	7.10	1.84	(1.29–3.28)	1.012 * / ÷ 1.052	6.83–7.55	0.898	(0.647 to 0.974)
<b>Trials 2–3</b>							
6-s distance (m)	37.59	1.33	(0.94–2.37)	1.002 * / ÷ 1.037	36.30–39.06	0.930	(0.748 to 0.982)
6-s mMP (W)	965.25	5.44	(3.87–10.01)	1.004 * / ÷ 1.158	836.84–1222.29	0.915	(0.700 to 0.978)
6-s mMSS ( $\text{m s}^{-1}$ )	7.33	1.32	(0.93–2.35)	1.007 * / ÷ 1.037	7.12–7.66	0.944	(0.793 to 0.986)
3-s distance (m)	18.27	1.83	(1.29–3.26)	1.001 * / ÷ 1.052	17.39–19.22	0.897	(0.642 to 0.973)
3-s mMP (W)	1043.75	10.05	(7.27–19.28)	1.036 * / ÷ 1.304	829.04–1409.58	0.793	(0.369 to 0.944)
3-s mMSS ( $\text{m s}^{-1}$ )	7.19	1.70	(1.20–3.03)	1.014 * / ÷ 1.048	6.96–7.64	0.915	(0.700 to 0.978)

CV: coefficient of variation; CI: confidence intervals; ICC: intraclass correlation coefficient; mMP: mean maximal power; and mMSS: mean maximal sprinting speed.

The reliability of 3-s and 6-s sprinting on the NMT in this study agrees with previous research.<sup>10</sup> Abt<sup>10</sup> reported a CV of 5.5% and 2.7% for SD and a CV of 4.1% and 3.5% for mMSS for a 3-s and 6-s sprint on the NMT, respectively. While these findings are larger when compared to the 3-s and 6-s sprint results of the present study (6 s SD CV = 1.33%, 6 s mMSS CV = 1.30%), they do support the finding of a greater reliability of a 6-s sprint compared to a 3-s sprint on the NMT. In addition, Abt<sup>10</sup> also reported a low CV of 2.2% for TD covered during a 15-min soccer-specific NMT protocol. This result is similar to the CV of 1.9% for TD covered during the team-sport simulation in this study. The consistent high levels of reliability of team-sport simulation distance measures shows that the team-sport simulation is a reliable tool for assessing performance demands of team-sport running activity.

To our knowledge this is the first study to extensively investigate the reliability of physiological measures during a NMT team-sport simulation. The mean ( $\pm$ SD) physiological responses of HR ( $160.0 \pm 9.7$  beats  $\text{min}^{-1}$ ),  $\dot{V}_{O_2}$  ( $3.0 \pm 0.3$  L  $\text{min}^{-1}$  or  $72.8\%$  of  $\dot{V}_{O_{2\max}} \pm 3.8\%$  of  $\dot{V}_{O_{2\max}}$ ) and post-exercise  $[\text{BLa}^-]$  ( $12.1 \pm 3.9$  mmol  $\text{L}^{-1}$ ) in this study were similar to those reported during field-based running protocols and other NMT soccer-specific simulations.<sup>3,5,10,24</sup> The present results show a high reproducibility of mean HR during the team-sport simulation with a TE of 2.6 beats  $\text{min}^{-1}$ . This is in agreement with previous research reporting a low TE for mean HR during the first (2.6 beats  $\text{min}^{-1}$ ) and second half (2.8 beats  $\text{min}^{-1}$ ) of a 90-min soccer-specific NMT protocol.<sup>3</sup> Similarly, low within-subject variability was shown for  $\dot{V}_{O_2}$  (CV = 5.3%) in the present study. While there are numerous studies researching the reliability of  $\dot{V}_{O_2}$  during steady state exercise, this current study is the first to assess the reliability of  $\dot{V}_{O_2}$  during non-steady state activity making comparisons difficult.<sup>25</sup> When comparing the  $\dot{V}_{O_2}$  during the team-sport simulation (CV = 5.3%) to steady state  $\dot{V}_{O_2}$  during 4-min intervals of running at 14 km  $\text{h}^{-1}$  (CV = 2.4%), 16 km  $\text{h}^{-1}$  (CV = 2.5%) and 18 km  $\text{h}^{-1}$  (CV = 2.4%), the variability in non-steady state exercise is much higher.<sup>25</sup> It should be noted that the reliability of the physiological responses are related to some degree to the reliability of the team-sport simulation test. That is, because the team-sport simulation test is 'self-chosen', the test in itself incorporates some random error, which in turn affects the variability of the physiological responses. Nonetheless, the CV for  $\dot{V}_{O_2}$  during the present study still remains under the criterion value of 10% to be considered reliable.<sup>22</sup>

**Table 3** Measures of reliability (with 95% confidence intervals) for physiological measures from three trials of a 30-min NMT team-sport running simulation ( $n = 11$ )

Physiological variable	Grand mean	CV (%)	CI (%)	Ratio limits of agreement	Lower-upper limit	ICC	CI
<b>Trials 1–2</b>							
$\dot{V}_{O_2}$ ( $L \cdot min^{-1}$ )	2.98	3.31	(2.34–5.97)	1.030 * / ÷ 1.094	2.81–3.36	0.750	(0.272 to 0.931)
Mean HR (beats $min^{-1}$ )	159.00	1.33	(0.93–2.36)	0.990 * / ÷ 1.037	151.69–163.22	0.955	(0.831 to 0.988)
Post [BLa <sup>-</sup> ] (mmol $L^{-1}$ )	12.30	24.45	(18.63–53.60)	1.124 * / ÷ 1.834	7.54–25.36	0.766	(0.308 to 0.936)
<b>Trials 2–3</b>							
$\dot{V}_{O_2}$ ( $L \cdot min^{-1}$ )	2.94	5.28	(3.76–9.71)	0.943 * / ÷ 1.153	2.40–3.20	0.598	(–0.003 to 0.882)
Mean HR (beats $min^{-1}$ )	158.00	1.71	(1.20–3.05)	0.993 * / ÷ 1.048	149.25–163.99	0.933	(0.758 to 0.983)
Post [BLa <sup>-</sup> ] (mmol $L^{-1}$ )	12.53	17.58	(13.07–36.15)	0.965 * / ÷ 1.567	7.72–18.94	0.648	(0.079 to 0.899)

CV: coefficient of variation; CI: confidence intervals; ICC: intraclass correlation coefficient; mean HR: mean heart rate; and post [BLa<sup>-</sup>]: post-exercise blood lactate concentration.

[BLa<sup>-</sup>] measured at the completion of the team-sport simulation is shown to have only moderate reliability (CV = 17.6%) and is above the criterion of 10% commonly used to determine a reliable measure.<sup>22</sup> Furthermore, when using the raw data (i.e., not log transformed) to calculate the TEM, the reliability value was 17.1%. This value is just above the acceptable range of 15%, which is a recommended standard for [BLa<sup>-</sup>] measurement.<sup>26</sup> However, this relatively high variability established in post-exercise [BLa<sup>-</sup>] in the present study is similar to the 17% CV reported for [BLa<sup>-</sup>] measured at the completion of the Yo-Yo Intermittent Recovery Test, a common field test used to assess team-sport performance.<sup>24</sup> On the basis of the present results, it appears that [BLa<sup>-</sup>] is not as reliable a measure as other physiological (i.e., HR and  $\dot{V}_{O_2}$ ) or performance measures. We recommend that [BLa<sup>-</sup>] should not be used as the sole physiological measure for assessing or monitoring team-sport running performance on the NMT.<sup>27</sup> Rather [BLa<sup>-</sup>] should be used in conjunction with either HR or  $\dot{V}_{O_2}$  and combined with key performance measures.

There is increasing awareness among researchers for the need for adequate familiarisation to minimise the learning effect associated with measuring tools.<sup>6</sup> The present study included one familiarisation session on the NMT, which incorporated the PSSA followed by one 15-min period of the team-sport simulation. However, it appears that this familiarisation session may not have been adequate in preparing subjects with the demands of NMT sprinting as demonstrated by the significant difference in sprinting variables and HIA between trials 1–2 and 1–3. It has been previously shown that NMT sprinting requires a greater forward body lean and a certain force (depending on mass) to be exerted to overcome the intrinsic resistance of the NMT.<sup>6</sup> In addition the exercise protocol developed in the present study was of a 'stop-start' nature, requiring frequent acceleration and deceleration (i.e., ~ every 10 s) to be applied to the NMT. Therefore, it may be that the subjects were not fully familiarised with the unique sprinting technique required on the NMT or with the intermittent nature of the protocol, causing slower reaction times and leading to differences in the amount of force and acceleration applied to a sprint. In contrast, the lack of any significant differences between trials 2 and 3 among physiological or performance variables shows that after two trials (i.e., familiarisation and trial 1) the subjects are fully familiarised with the ergonomics of NMT sprinting. This is in agreement with previous research reporting a negligible learning effect in consecutive trials

of short duration, high-intensity cycling after two familiarisation sessions.<sup>28</sup> We suggest that future studies investigating NMT team-sport running use two familiarisation trials to ensure that learning effects are negligible. These familiarisation trials should include the testing protocol and concentrate on NMT sprinting technique, assuring that acceleration and deceleration have been mastered prior to testing.

On the basis of the present results, we suggest that a minimum sample size of between 5 and 9 is adequate to detect a 10% change in the majority of physiological and performance variables measured during the team-sport simulation.<sup>22</sup> However, the minimum sample size needed for detecting a 10% change in 3 s mMSS and post-exercise [BLa<sup>-</sup>] is much larger, being 25 and 86, respectively. In order to detect a smaller change such as 5%, minimum sample sizes are larger again ranging from between 8 and 28 for the majority of the physiological and performance variables and 100 and >200 for 3-s mMSS and post-exercise [BLa<sup>-</sup>], respectively. These latter sample size estimations of 100 and >200 would be unfeasible in most research study designs. Similarly, in agreement with previous research,<sup>8</sup> the extremely large sample sizes needed to detect small changes (i.e.,  $\geq 1\%$ ) which prove meaningful for elite athletes<sup>29</sup> would clearly be impractical in NMT team-sport simulation studies.

### Practical implications

- A non-motorised treadmill can be used to assess and monitor changes in both physiological and performance measures across team sports.
- Two practice sessions focusing on correct non-motorised treadmill sprinting technique and acceleration during changes in speed should be completed before testing to increase the reliability.
- A 6-s sprint is best for assessing non-motorised treadmill sprint performance in team-sport athletes.

### Acknowledgements

Dr. Grant Abt assisted in the development of the NMT protocol. Funding was through an Early Career Research Grant from the University of Technology, Sydney.

### References

1. Douge B. Football: the common threads between the games. In: *Science and football*. London: E & FN Spon; 1988. p. 3–19.
2. Reilly T, Gilbourne D. Science and football: a review of applied research in the football codes. *J Sports Sci* 2003;21(9):693–705.
3. Thatcher R, Batterham AM. Development and validation of a sport-specific exercise protocol for elite youth soccer players. *J Sports Med Phys Fit* 2004;44(1):15–22.
4. Sirotic AC, Coutts AJ. Physiological and performance test correlates of prolonged, high-intensity, intermittent running performance in moderately trained women team sport athletes. *J Strength Cond Res* 2007;21(1):138–44.
5. Abt GA, Reaburn PRJ, Holmes MA, Gear T. Changes in peak speed during prolonged high-intensity exercise that simulates team sport play. *J Sports Sci* 2003;21(4):256–7.
6. Lakomy HKA. The use of a non-motorised treadmill for analysing sprint performance. *Ergonomics* 1987;30(4):627–37.
7. Coutts AJ, Reaburn PRJ, Abt GA. Heart rate, blood lactate concentration and estimated energy expenditure in a semi-professional rugby league team during a match: a case study. *J Sports Sci* 2003;21(2):97–103.
8. Tong RJ, Bell W, Ball G, Winter EM. Reliability of power output measurements during repeated treadmill sprinting in rugby players. *J Sports Sci* 2001;19(4):289–97.
9. Hughes MG, Doherty M, Tong RJ, Reilly T, Cable NT. Reliability of repeated sprint exercise in non-motorised treadmill ergometry. *Int J Sports Med* 2006;27(11):900–4.
10. Abt GA. *An attempt to assess the effect of oral creatine monohydrate supplementation on prolonged, high-intensity, intermittent exercise that simulates team sport play*. Unpublished doctoral thesis. North Rockhampton, Queensland: School of Health and Human Performance, Faculty of Arts, Health and Sciences, Central Queensland University. 356 pp.
11. Atkinson G, Nevill AM. Statistical methods for assessing measurement error (reliability) in variables relevant to sports medicine. *Sports Med* 1998;26(4):217–38.
12. Coutts AJ, Wallace LK, Slattery KM. Monitoring changes in performance, physiology, biochemistry, and psychology during overreaching and recovery in triathletes. *Int J Sports Med* 2007;28(2):125–34.
13. Bangsbo J, Nørregaard L, Thorsøe F. Activity profile of competition soccer. *Can J Sport Sci* 1991;16(2):110–6.
14. Duthie G, Pyne D, Hooper S. Applied physiology and game analysis of rugby union. *Sports Med* 2003;33(13):973–91.
15. Meir R, Arthur D, Forrest M. Time and motion analysis of professional rugby league: a case study. *Strength Condition Coach* 1993;3(1):24–9.
16. Appleby B, Dawson B. Video analysis of selected game activities in Australian Rules football. *J Sci Med Sport* 2002;5:129–42.
17. Keane S, Reilly T, Hughes M. Analysis of work-rates in Gaelic football. *Aust J Sci Med Sport* 1993;25(4):100–2.
18. Spencer M, Lawrence S, Rechichi C, Bishop D, Dawson B, Goodman C. Time–motion analysis of elite field hockey, with special reference to repeated-sprint ability. *J Sports Sci* 2004;22(9):843–50.
19. Mohr M, Krustup P, Bangsbo J. Match performance of high-standard soccer players with special reference to development of fatigue. *J Sports Sci* 2003;21(7):519–28.

Please cite this article in press as: Sirotic AC, Coutts AJ, The reliability of physiological and performance measures during simulated team-sport running on a non-motorised treadmill, *J Sci Med Sport* (2007), doi:10.1016/j.jsams.2007.04.008

20. Hopkins WG. *Reliability from consecutive pairs of trials (Excel spreadsheet)*. Available at: [sportssci.org/resource/stats/xrely.xls](http://sportssci.org/resource/stats/xrely.xls) [accessed June 22, 2006].
21. Bland JM, Altman DG. Statistical method for assessing agreement between two methods of clinical measurement. *Lancet* 1986;1:307–10.
22. Atkinson G, Nevill AM, Edwards B. What is an acceptable amount of measurement error? The application of meaningful 'analytical goals' to the reliability of sports science measurements made on a ratio scale. *J Sports Sci* 1999;17:18.
23. Vincent WJ. *Statistics in kinesiology*. 2nd ed. Lower Mitcham: Human Kinetics; 1999.
24. Krustup P, Mohr M, Amstrup T, Rysgaard T, Johansen J, Steensberg A, et al. The Yo-Yo intermittent recovery test: physiological response, reliability, and validity. *Med Sci Sports Exerc* 2003;35(4):697–705.
25. Saunders PU, Pyne DB, Telford RD, Hawley JA. Reliability and variability of running economy in elite distance runners. *Med Sci Sports Exerc* 2004;36(11):1972–6.
26. Gore CJ. Quality assurance in exercise physiology laboratories. In: *Physiological tests for elite athletes*. 1st ed. Champaign, IL: Human Kinetics; 2000. p. 3–11.
27. Krustup P, Mohr M, Steensberg A, Bencke J, Kjaer M, Bangsbo J. Muscle and blood metabolites during a soccer game: implications for sprint performance. *Med Sci Sports Exerc* 2006;38(6):1165–74.
28. Capriotti PV, Sherman WM, Lamb DR. Reliability of power output during intermittent high-intensity cycling. *Med Sci Sports Exerc* 1999;31(6):913–5.
29. Hopkins WG. Measures of reliability in sports medicine and science. *Sports Med* 2000;30(1):1–15.

Available online at [www.sciencedirect.com](http://www.sciencedirect.com)

