SESNZ
SPORT & EXERCISE SCIENCE
NEW ZEALAND

The Journal of Sport and Exercise Science, Vol. 10, Issue 1, 45-53 (2026)

www.jses.net

JSES
ISSN: 2703-240X

Validity and reliability of England Hockey's 16-pitch length aerobic capacity test

Thomas Spooner[1*], Paul Wallis[1], Andrew West[1]

[1]University of Chichester, School of Sport, Science and Engineering, United Kingdom

ARTICLE INFO

Received: 25.09.2025
Accepted: 13.02.2026
Online: 13.07.2026

Keywords:
Maximal aerobic speed
Validity
Reliability
Exercise physiology
Team sport
Field hockey

ABSTRACT

Maximal aerobic speed (MAS) is the minimum speed required to achieve maximal oxygen uptake ($\dot{V}O_{2max}$) and serves as a practical tool for prescribing training intensities and assessing aerobic capacity. The 16-pitch length time trial test (TT_{16PLT}) is commonly used within the England Hockey pathway; however, there is currently no data pertaining to the validity and reliability of the TT_{16PLT} to predict MAS. The present study aimed to determine the validity and reliability of the TT_{16PLT} to measure MAS. Twenty well-trained male field hockey players (mean ± SD: age 22.7 ± 3.8 y, stature 1.82 ± 0.09 m, mass 79.6 ± 11.2 kg, and $\dot{V}O_{2max}$ 52.6 ± 5.6 mL·kg^{-1}·min^{-1}) participated, completing the TT_{16PLT} twice and a treadmill assessment to establish laboratory derived MAS. Results reported no significant difference in MAS between repeated TT_{16PLT} trials (4.02 ± 0.30 vs 4.01 ± 0.30 m·s^{-1}, $p > 0.05$), narrow 95% limits of agreement (LoA: ± 0.11 m·s^{-1}), negligible systematic bias (0.01 m·s^{-1}), excellent reliability (ICC = 0.98) and low variability (CV = 1.7%). The LoA indicate an overestimation in TT_{16PLT} measured MAS compared to laboratory derived MAS (4.04 ± 0.30 vs 3.96 ± 0.38 m·s^{-1}, $p > 0.05$), with wide LoA (± 0.59 m·s^{-1}). TT_{16PLT} is a reliable measure of MAS; however, it can be concluded that the TT_{16PLT} derived MAS may not provide sufficient validity for precise exercise prescription.

1. Introduction

Field hockey is a widely played team invasion sport (Lord et al., 2022), with over 800 clubs affiliated with the national governing body, England Hockey, and approximately 140,000 individuals participating regularly within the club system (England Hockey, 2026). It is classified as 'heavy exercise' due to its high physiological demands (Reilly & Borrie, 1992), requiring players to sustain prolonged periods of high intensity intermittent activity characterised by repeated accelerations, decelerations, and changes of direction, thereby demanding substantial aerobic energy contribution alongside frequent maximal anaerobic efforts (Lombard et al., 2021). Consequently, field hockey players are highly conditioned athletes, exhibiting aerobic power values of 48-65 mL·kg^{-1}·min^{-1}, with elite level players attaining a maximal aerobic uptake ($\dot{V}O_{2max}$) in excess of 60 mL·kg^{-1}·min^{-1} (Kostiukevych et al., 2021; Reilly & Borrie, 1992). Supporting these physiological demands, National League Hockey (United Kingdom) players on average cover a total distance of 5986 ± 1105 m per 70-min match, at a playing intensity of 116 ± 12 m.min^{-1} (Noblett et al., 2023).

Maximal aerobic speed (MAS) is the minimum speed required for an individual to achieve $\dot{V}O_{2max}$ (Bellenger et al., 2015). Research would suggest that training at or above 100% MAS is necessary for improving aerobic power (Baquet et al., 2001; Dupont et al., 2002; Tabata et al., 1996). Nevertheless, the main benefit of completing high-intensity intermittent training is that it is more time-efficient and logistically easier to implement into the busy training and competition schedules of team sport athletes (Baker, 2011), while still generating adequate stimuli to induce similar aerobic adaptations to traditional moderate-intensity endurance training but with the addition of anaerobic benefits (Tabata et al., 1996). Similar findings have been observed when completing supramaximal intervals at 120% MAS, which resulted in a significantly longer duration spent exercising at $\dot{V}O_{2max}$ compared to continuous work at 100% MAS (Dupont et al., 2002).

*Corresponding Author: Thomas Spooner, School of Sport, Science and Engineering, University of Chichester, United Kingdom, t.spooner@chi.ac.uk

Increasing the time at $\dot{V}O_{2max}$ causes greater metabolic stress and consequently induces further aerobic adaptations (Myrkos et al., 2022), hence why the duration of training spent at or above $\dot{V}O_{2max}$ is vital to improving aerobic power (Baquet et al., 2001; Dupont et al., 2002; Tabata et al., 1996). Baquet et al. (2001) reported that completing one hour a week of high-intensity short intermittent exercises (10-s) at 100-120% can have a significant improvement on standing broad jump, 20-m shuttle run and 7-min time trial (TT) performance in ten weeks of training. Interestingly, this study highlights that not only can high-intensity intermittent training improve aerobic power, but it can also have significant benefit on explosive power which is an essential component of success in field sports, including field hockey (Elferink-Gemser et al., 2006). There are several strategies available to implement MAS prescribed high-intensity intermittent exercise into training, such as the 'Eurofit' method (Dupont et al., 2004; Wong et al., 2010), 'Tabata' method (Tabata et al., 1996), and maximal aerobic grids (Baker, 2011).

A high level of aerobic capacity has been identified as essential for success in field hockey (Reilly & Borrie, 1992), henceforth coaching staff regularly monitor players aerobic capacity. Laboratory testing, using indirect calorimetry to obtain $\dot{V}O_{2max}$ while running on a treadmill, is considered the 'gold standard' measure of MAS (Baker & Heaney, 2015). However, laboratory testing is costly, time-consuming, and impractical for team settings; therefore, MAS is often calculated using field tests as a more practical and time effective indicator of athletes' aerobic capacity (Baker & Heaney, 2015). Commonly used field tests such as the 20 m Multi-Stage Shuttle Test (MSST) and the Yo-Yo Intermittent Recovery Test (Yo-Yo IRT) provide valid and reliable estimates of $\dot{V}O_{2max}$ (Krustrup et al., 2003; Leger & Lambert, 1982) but are less suitable for prescribing training loads because results are expressed in stages rather than speed (Dupont et al., 2010). In contrast, TT protocols lasting between five and seven minutes offer a practical and reliable method for determining MAS, producing values strongly correlated with laboratory measures (Baker & Heaney, 2015). For example, MAS calculated from a 5-min TT was strongly correlated with the MAS established during a graded treadmill exercise test using indirect calorimetry ($r = 0.94$, $p < 0.001$) (Berthon et al., 1997). Similarly, MAS calculated from the mean speed during a 1.2km shuttle run test strongly correlate with the MAS determined in the 30–15 Intermittent Fitness Test (30-15 IFT), where MAS is defined as the velocity attained during the last completed stage (Kelly & Wood, 2013). Furthermore, the 1.2km shuttle run test exhibited strong repeatability with an intraclass correlation coefficient (ICC) of 0.99 for performance time and MAS (Brew & Kelly, 2014). Interestingly, all reported performance times were between five and seven minutes (Kelly et al., 2014), reinforcing this duration as optimal for accurately establishing MAS during a TT (Baker & Heaney, 2015). Furthermore, Bellenger et al., (2015) reported similar findings which were that MAS can be predicted from the average speed during a TT for any distance between 1200 m and 2200 m. Interestingly, the average speed established during the 2000 m TT had the narrowest limits of agreement and smallest bias compared with pre-established MAS (Bellenger et al., 2015), indicating that 2000 m TT may be the optimal distance to predict MAS. Collectively, TT based MAS testing provide a simple and time effective solution for determining athletes' aerobic capacity.

The 16-pitch length TT test (TT_{16PLT}) is routinely used within the England Hockey pathway as a performance measure to track athletes' aerobic capacity and fitness (England Hockey Athlete Loading, 2025). Despite its widespread application in field hockey, there is currently no published evidence regarding the validity and reliability of the TT_{16PLT} for predicting MAS. In theory, the TT_{16PLT} should provide an accurate estimate of MAS, as the total distance of 16 pitch lengths (1462.4 m) aligns with time-trial distances previously shown to be suitable for MAS prediction (Bellenger et al., 2015). Additionally, well-trained hockey players would be expected to complete this distance within the desired five to seven minutes required to establish MAS (Baker & Heaney, 2015). However, in the absence of empirical evidence, the suitability of the TT_{16PLT} for accurately measuring MAS cannot currently be evaluated. Therefore, the present study aims to determine the validity and reliability of the TT_{16PLT} for measuring MAS compared to laboratory 'gold standard' derived MAS values in well-trained male hockey players.

2. Methods

2.1. Participants

Twenty well-trained male field hockey players (mean ± SD: age 22.7 ± 3.8 y, stature 1.82 ± 0.09 m, mass 79.6 ± 11.2 kg, and $\dot{V}O_{2max}$ 52.6 ± 5.6 $mL \cdot kg^{-1} \cdot min^{-1}$) playing National League Hockey in the Men's Conference West (United Kingdom) volunteered to participate in this study. All participants provided written informed consent and completed a health screening assessment once the experimental procedures and the associated risks of participation were explained. Ethical approval was obtained according to regulations by the University of Chichester's Research Ethics Committee for research involving the use of human participants.

2.2. Preliminary testing

All participants were familiar with the TT_{16PLT} having routinely completed the test as part of a conditioning programme. During preliminary testing participants were reminded the TT_{16PLT} should be completed with maximal effort and were instructed to cover the distance in the quickest time possible. All participants completed the total distance (1462.4 m) in an acceptable time (5-7 mins) to accurately calculate MAS (Baker & Heaney, 2015).

2.3. Experimental overview

Participants reported to the University of Chichester on three separate occasions, twice to complete the TT_{16PLT} and a third visit to complete a submaximal and maximal treadmill test in the laboratory. There was no more than a week and a minimum of 72 hours between each testing session for all participants. Participants were required to maintain their normal diet and exercise volume throughout the course of the study. Additionally, they were instructed to abstain from alcohol, strenuous exercise, and caffeine 24 hours prior to any testing. All participants wore standard field hockey attire, including field hockey specific shoes, to each testing session. Additionally, participants completed both their first and second TT_{16PLT} at the same time of day (± 1 hour), with low wind speeds (< 5 km·h^{-1}), and in dry conditions (temperature ~18°C, humidity ~40%) to control for environmental variations (Kusumoto et al., 2021) and changes in ambient temperature (No & Kwak, 2016). Environmental conditions were measured using a handheld anemometer (Kestrel 5500 Pocket Weather Meter, Kestrel, Boothwyn, USA). All TT_{16PLT} were conducted on the University of Chichester AstroTurf pitch; an International Hockey Federation regulation sized field hockey pitch, rectangular field of play 91.4 m long and 55.0 m wide (Fédération Internationale de Hockey, 2023). Furthermore, laboratory conditions were kept consistent (temperature 18°C, humidity 40%) between participants, using the same equipment and apparatus for each testing session.

2.4. Procedure: 16-pitch length time trial test

Participants completed a standardised warm-up; consisting of a 5-min self-paced run around the pitch perimeter, followed by a dynamic stretching protocol (Perrier et al., 2011). Participants started on the baseline and were instructed to run to the opposite baseline and back eight times, equating to a total of 16-pitch lengths completed and a total distance of 1462.4 m covered. Participants were required to touch the baseline with their foot before they could turn to run in the opposite direction, standardised verbal encouragement was provided to ensure all participants were giving maximal effort and crossed the finish line in the quickest time possible. The TT_{16PLT} was conducted in groups of at least three participants; however, group sizes never exceeded twelve participants. MAS was then calculated using the equation MAS (m·s^{-1}) = total distance (m) / time taken (s).

2.5. Procedure: Laboratory testing

Upon arrival, body mass was recorded using a scale (Seca Model 880, Seca Ltd Birmingham, UK) and height was measured using a stadiometer (Seca 213 Portable Stadiometer, Seca Medical Scales, Birmingham, UK). Capillary blood samples were taken from the participants left index finger and analysed using an automated analyser (Biosen C-Line; EKF Diagnostic, Ebendorfer Chaussee 3, Barleben, Germany) to establish baseline blood lactate values. Participants were then fitted with a heart rate (HR) monitor and chest strap (V800, Polar Electro UK Ltd, Warwick, UK). Both the submaximal and maximal treadmill protocols were conducted on a HP Cosmos treadmill (HP Cosmos Pulsar, h/p/cosmos sports & medical gmbh, Nussdorf-Traunstein, Germany). Breath-by-breath oxygen consumption ($\dot{V}O_2$) measures were recorded using a metabolic analyser system (Cosmed Quark CPET, Cosmed SRL, Rome, Italy), which was calibrated according to the manufacturer's instructions. A bidirectional digital turbine measured inspired and expired airflow, whereas an electrochemical cell O_2 analyser and non-dispersive infrared CO_2 analyser simultaneously measured expired gases. Participants wore an oronasal facemask (Vmask, Hans Rudolph, Kansas City, USA) with a low dead space that was securely attached to the turbine. The inspired and expired gas volume and gas concentration signals were continuously sampled via a capillary line connected to the mouthpiece. The gas analysers were calibrated before each test with gases of known concentration, and the turbine was calibrated using a 3-liter syringe (Hans Rudolph, Kansas City, USA).

Participants all completed a warm-up on the treadmill at a self-selected speed for 10-mins, followed by personal stretches to replicate the stretches that would usually be completed prior to training or competition. The initial starting speed of the treadmill for the submaximal protocol was set at 60% of the average speed (m·s^{-1}) each participant achieved during the TT_{16PLT}. The treadmill gradient was set at 1% to reflect the energetic cost of outdoor running, accounting for air resistance (Jones & Doust, 1996). The submaximal ramp protocol was programmed so that every 4-mins the treadmill increased in speed by 0.28 m·s^{-1} (Galbraith, 2021). During the last minute of each stage, a blood lactate sample was collected and analysed. Once the participants blood lactate exceeded 4.0 mmol·L^{-1} of lactate concentration the exercise protocol was terminated (Galbraith, 2021), and the participant was given 15-mins of recovery. All participants reached 4.0 mmol·L^{-1} of lactate within 4-6 stages, providing adequate data to plot the relationship between running speed and $\dot{V}O_2$. After the 15-mins of recovery the maximal protocol was conducted, the starting speed was set at the same speed as for the submaximal protocol. However, the ramp protocol increased at a rate of 0.28 m·s^{-1} every minute until the participant reached volitional exhaustion (Galbraith, 2021). Post blood lactate values were recorded, 1-min, 3-min, and 5-min after exercise had ceased, to identify a peak lactate value. All participants gave a maximal effort, satisfying at least two of the three criteria upon termination of the test; HR within 10 beats·min^{-1} of age predicted maximum (Tanaka et al., 2001), an RER ≥ 1.14 (Howley et al., 1995), and signs of visible exhaustion (Midgley et al., 2017). The breath-by-breath data collected during both the submaximal and maximal test were exported to Excel (Microsoft Excel, V.16.70). Any irregular breaths were removed, and the data was smoothed using a rolling 60-s average. Participants maximum heart rate (HR_{max}) and $\dot{V}O_{2max}$ was identified during the maximal test by the highest HR and $\dot{V}O_2$ value recorded. Using the submaximal test data, $\dot{V}O_2$ at each incremental speed was identified and plotted on a scatterplot to establish a running speed and $\dot{V}O_2$ relationship. Linear regression was used to calculate the gradient of the line ($y = mx + c$) and as participants $\dot{V}O_{2max}$ was a known value from the maximal test the equation could be rearranged to calculate the running speed at which $\dot{V}O_{2max}$ occurred in m·s^{-1} (MAS).

2.6. Statistical approach

Visual checks for normal distribution, including histograms, boxplots, and Q-Q plots; all suggested that data was normally distributed with the absence of outliers. All data met the $z_{skewness}$ and $z_{kurtosis}$ +2 to -2 criteria for normality (O'Donoghue, 2013). Absolute agreement between TT_{16PLT} MAS vs laboratory MAS, and trial 1 vs trial 2 (TT_{16PLT}) were determined by the limits of agreement (LoA) analysis and Bland-Altman plots (Altman & Bland, 1983). LoA were calculated using the equation ± 1.96 multiplied by the standard deviation (SD) of the differences (Bland & Altman, 1986). Atkinson and Nevill (1998) recommended that a sample size greater than 40 is required to ensure 95% LoA can be applied to the wider population. However, the present study only had a sample size of 20; therefore, to indicate the precision of systematic bias and 95% LoA, the 95%

confidence intervals (CI) were derived from the calculation of standard error. Typical error of measure was calculated from the SD of the mean difference using the formula typical error of measurement = $SD_{(diff)}/\sqrt{2}$ (Hopkins, 2000). Expressed as a mean coefficient of variation (CV %) and the 95% CI were determined for CV using the methods described by Hopkins (2007). ICC (two way random, single measure) were calculated along with their 95% CI to facilitate comparisons with previous literature. According to Koo and Li (2016), the magnitude of the correlation was considered to be poor (ICC < 0.5), moderate (ICC, between 0.5 and 0.7), good (ICC, between 0.7 and 0.9), and excellent (ICC > 0.9). Paired sample t-test was used to determine differences between means. Cohen's *d* was used as effect size statistics (Cohen, 2013). Descriptive statistics are reported as mean and SD. Statistical significance was set at $\alpha = 0.05$ (Hinton, 2014). All analysis was completed in JASP V10.

3. Results

All twenty participants successfully completed the study. Physiological response variables to the maximal treadmill test, including $\dot{V}O_{2max}$, HR_{max}, respiratory exchange ratio (RER) and post-maximal blood lactate concentration are presented in Table 1. Peak treadmill speed ($m{\cdot}s^{-1}$) achieved during the maximal treadmill test is also displayed in Table 1.

Table 1. Physiological responses ($n = 20$) to the maximal treadmill test (mean ± SD).

Variables	*M* ± SD
$\dot{V}O_{2max}$ ($mL{\cdot}kg^{-1}{\cdot}min^{-1}$)	52.6 ± 5.6
HR_{max} ($beats{\cdot}min^{-1}$)	200 ± 9
RER	1.17 ± 0.05
Post-lactate ($mmol{\cdot}L^{-1}$)	11.0 ± 2.3
Peak Treadmill Speed ($m{\cdot}s^{-1}$)	4.81 ± 0.43

Notes: HR_{max} = Maximum Heart Rate; RER = Respiratory Exchange Ratio; $\dot{V}O_{2max}$ = Maximal Oxygen Uptake.

As illustrated in Table 2 and Figure 1, there is extremely good agreement between Trial 1 and Trial 2 (TT_{16PLT}) in measuring MAS. The mean difference (bias) is very close to zero (0.01 $m{\cdot}s^{-1}$), indicating negligible systematic bias between the two trials. Furthermore, the LoA 95% CI are very narrow, highlighting there is marginal difference between both trials for all participants. The CV of 1.7% suggests very low variability in the MAS measurements between Trial 1 and Trial 2. Additionally, the excellent ICC value (0.98) indicates there is high agreement in the MAS values obtained in both the trials. Moreover, there is no significant difference between the means of both trials ($t_{19} = 0.812$, $p = 0.812$, $p > 0.05$, $d = -0.182$).

Table 2. The reliability of the TT_{16PLT} ($n = 20$).

Measure	Mean ± SD		CV % [95 % CI]	Limits of Agreement			ICC [95 % CI]
	Trial 1	Trial 2		Bias [95 % CI]	Lower Limit [95 % CI]	Upper Limit [95 % CI]	
MAS ($m{\cdot}s^{-1}$)	4.02 ± 0.30	4.01 ± 0.30	1.7 [0.9 to 2.5]	0.01 [-0.02, 0.04]	-0.10 [-0.15, -0.06]	0.12 [0.08, 0.17]	0.98 [0.96, 0.99]

Notes: CI = Confidence Interval; CV = Coefficient of Variation; ICC = Intraclass Correlation Coefficient; MAS = Maximal Aerobic Speed; SD = Standard Deviation.

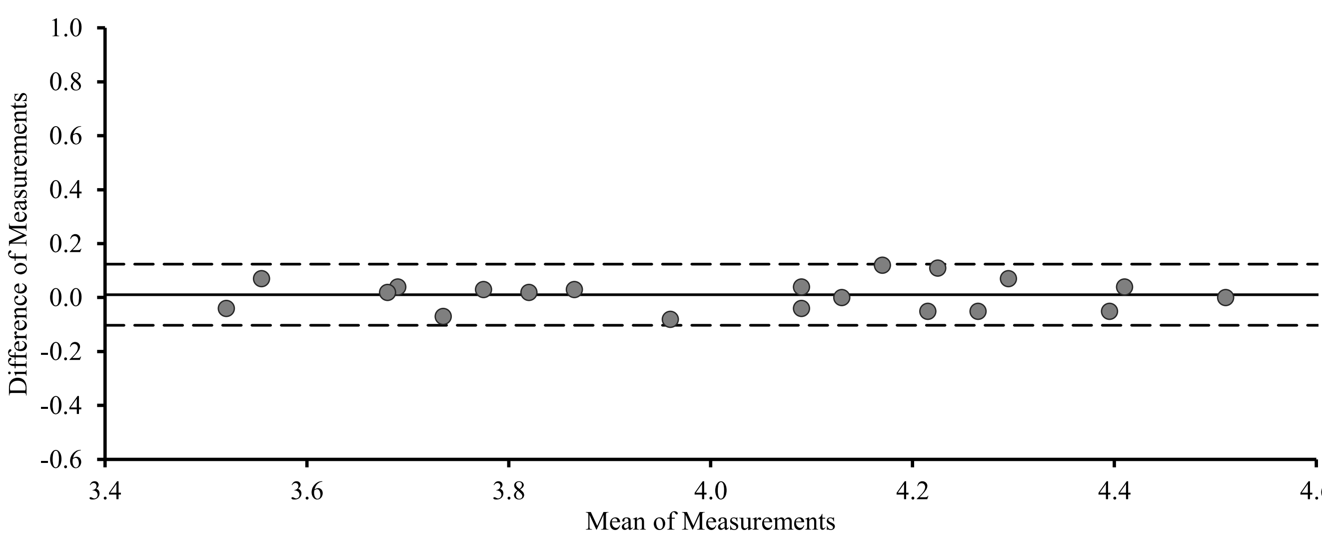

Figure 1. Bland and Altman plot (95% CI) of Trial 1 versus Trail 2 (TT_{16PLT}) MAS values ($n = 20$). The solid line represents the bias. The dotted lines above and below the centre solid line represent the upper and lower limits of agreement, respectively. *Notes:* Axis in Figure 1 and 2 have been set to the same scale to allow for an easier comparison of the LoA between Trial 1 vs Trial 2 and Field vs Laboratory.

As demonstrated in Table 3 and Figure 2, the MAS established from the TT_{16PLT} overestimates MAS compared to the laboratory test, with a 0.08 m·s^{-1} mean bias between the TT_{16PLT} and laboratory test, where 95% of the differences fall within a range of -0.05 to 0.22 m·s^{-1}. The LoA 95% CI are relatively wide, indicating a lack of precision and larger margin of error that individuals TT_{16PLT} MAS will not closely correlate with laboratory measured MAS. However, the ICC (0.62) does suggest that there is moderate agreement between the two methods of measuring MAS, with a CV of 8.5%. Paired sample t-test indicates that there is no statistically significant difference between the means of both methods at measuring MAS (t_{19} = -1.286, $p = 0.214$, $p > 0.05$, $d = -0.288$).

Table 3. The validity of the TT_{16PLT} (Field) compared to 'gold standard' Laboratory derived MAS (n = 20).

Measure	Mean ± SD		CV % [95 % CI]	Limits of Agreement			ICC [95 % CI]
	Field	Laboratory		Bias [95 % CI]	Lower Limit [95 % CI]	Upper Limit [95 % CI]	
MAS (m·s^{-1})	4.04 ± 0.30	3.96 ± 0.38	8.5 [4.0 to 13.1]	0.08 [-0.05 to 0.22]	-0.50 [-0.74 to -0.26]	0.67 [0.43 to 0.91]	0.62 [0.27 to 0.83]

Notes: CI = Confidence Interval; CV = Coefficient of Variation; ICC = Intraclass Correlation Coefficient; MAS = Maximal Aerobic Speed; SD = Standard Deviation.

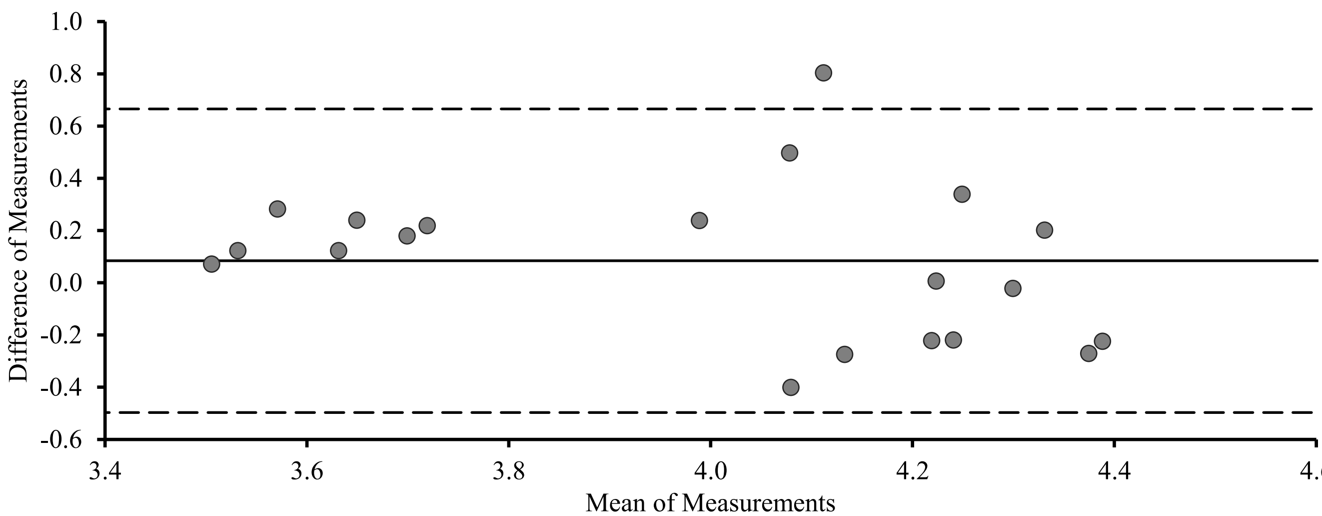

Figure 2. Bland and Altman plot (95% CI) of participants greatest TT_{16PLT} MAS versus laboratory measured MAS (n = 20). The solid line represents the bias. The dotted lines above and below the centre solid line represent the upper and lower limits of agreement, respectively.

Table 4 shows that if the TT_{16PLT} is used to prescribe exercise intensity (running speed) that there is a substantial disparity between the actual percentage of MAS the individual is working at and the prescribed MAS. For example, if the prescribed MAS was 120% (4.80 m·s^{-1}) then there would be confidence that at least 95% of the population would be exercising between 108% and 137% of their 'true' MAS. The magnitude of error, 12% underestimation or 13% overestimation of 'true' MAS, is consistent regardless of the prescribed MAS.

Table 4. Potential variation in running speed (m·s^{-1}) and prescribed MAS (%) in relation to the LoA lower limit (-0.50 m·s^{-}1) and upper limit (0.67 m·s^{-1}) 95% CI (Example for an individual with a 'true' MAS of 4.00 m·s^{-1}).

Prescribed MAS (%)	Target Running Speed (m·s^{-1})	Actual Running Speed (m·s^{-1})		Actual MAS (%)	
		Lower Bound [95 % LoA]	Upper Bound [95 % LoA]	Lower Bound [95 % LoA]	Upper Bound [95 % LoA]
90%	3.60	3.10	4.27	78%	107%
100%	4.00	3.50	4.67	88%	117%
120%	4.80	4.30	5.47	108%	137%
140%	5.60	5.10	6.27	128%	157%

Notes: LoA = Limits of Agreement; MAS = Maximal Aerobic Speed.

4. Discussion

The present study aimed to determine the validity and reliability of the TT_{16PLT} test in measuring MAS compared to laboratory 'gold standard' derived MAS values in well-trained male hockey players. The primary finding of the study was that the TT_{16PLT} has high test-retest reliability, with negligible systematic bias between trials. Another key finding was that the TT_{16PLT} provided an overestimation of MAS compared to laboratory calculated MAS, with a 0.08 $m{\cdot}s^{-1}$ mean difference between the two measurements.

The TT_{16PLT} demonstrated excellent agreement in MAS between trials (ICC = 0.98) with a mean bias close to zero (0.01 $m{\cdot}s^{-1}$). The narrow LoA 95% CI (Figure 1) and low variability (CV = 1.7%) indicate minimal variability in repeated measures (Table 2). No significant difference was found between trial means ($p > 0.05$). Overall, the TT_{16PLT} shows very high test-retest reliability, reporting ICC and CV values comparable to well-established field tests commonly used to assess team sport athletes' aerobic fitness (Aziz et al., 2005; Grgic et al., 2021; Thomas et al., 2006). The 1.2km shuttle run test (Kelly & Wood, 2013) that was developed to measure MAS has shown excellent repeatability, with an ICC of 0.99 for performance time and MAS (Brew & Kelly, 2014). Similarly, the 30-15 IFT demonstrates high test-retest reliability, with ICCs ranging from 0.80 to 0.99 across seven studies in a systematic review (Grgic et al., 2021), and the Yo-Yo IRT has proven reliable in both recreational and elite athletes (Thomas et al., 2006). The MSST also reports strong reliability, showing no significant difference ($p > 0.05$) in predicted $\dot{V}O_{2max}$ between repeated MSST trials in both endurance and team sport athletes, with an ICC of 0.90 and CV of 3.5%. In comparison, the TT_{16PLT} achieved excellent agreement (ICC = 0.98) and low variability (CV = 1.7%) between trials, indicating test-retest reliability comparable to these established field tests and supporting its use for tracking changes in physical performance over time.

The present study reported that the TT_{16PLT} derived MAS has moderate agreement with MAS obtained by the laboratory 'gold standard' test (ICC = 0.62). However, the TT_{16PLT} appears to overestimate MAS compared to the laboratory test (Figure 2). A 0.08 $m{\cdot}s^{-1}$ mean bias between the two methods was observed, with 95% of the differences falling within a range of -0.05 to 0.22 $m{\cdot}s^{-1}$ (Table 3). The relatively wide LoA 95% CI (Figure 2) and CV of 8.5% (Table 3) would suggest that the TT_{16PLT} measurement of MAS is lacking in precision and has high variability between individuals. Therefore, it would be advised not to use the TT_{16PLT} to measure MAS for research purposes, as the predictive value will likely be variable from the 'true value'. In a similar study, Thron et al. (2022) reported that a 1500 m TT showed systematic overestimation of MAS. The overestimation in both studies could be a result of the TT distance being too short for the primary energy system to be aerobic, thereby requiring a substantial anaerobic energy contribution to achieve the fastest possible completion time (Thron et al., 2022). Bellenger et al. (2015) reported that the average speed established during the 2000 m TT had the narrowest limits of agreement and smallest bias, potentially indicating that a longer TT distance reduces the anaerobic contribution to energy production resulting in a more precise MAS measurement.

The wide LoA 95% CI, lower limit (-0.50 $m{\cdot}s^{-1}$) and upper limit (0.67 $m{\cdot}s^{-1}$), compromise the ability to accurately prescribe exercise intensities from the MAS established in the TT_{16PLT}, as seen in Table 4. Overestimating the prescription of intermittent training intensities can potentially be detrimental to exercise performance and health, as it might result in individuals accidentally overtraining, increasing the risk of injury and fatigue (Kellmann, 2010). On the other hand, if the prescribed exercise intensity is an underestimation, then the amount of metabolic stress the individual is exposed to may be inadequate to induce sufficient training adaptations (Myrkos et al., 2022). The duration of training spent at $\dot{V}O_{2max}$ has been shown to be essential in improving aerobic power (Baquet et al., 2001; Dupont et al., 2002; Tabata et al., 1996). Dupont et al. (2002) reported that participants spent a significantly longer duration at $\dot{V}O_{2max}$ ($p > 0.05$) when completing intermittent running bouts at 120% of MAS, compared to intermittent runs at 110%, 130% and 140% of MAS. Interestingly, intermittent running bouts at 130% and 140% of MAS had the shortest duration spent exercising at $\dot{V}O_{2max}$ (Dupont et al., 2002). This information would suggest for a team sport athlete looking to improve their aerobic power, they should be aiming to complete high-intensity intermittent exercise sessions at 120% of their MAS (Dupont et al., 2002). However, if they calculated their MAS using the TT_{16PLT} the actual percentage of MAS that they would likely be working at would be anywhere between 108% and 137% (Table 4). The uncertainty of the exact percentage of MAS that they are exercising at makes the TT_{16PLT} impractical for the prescription of exercise intensities. Additionally, exercising at 137% of MAS would result in a shorter duration spent at $\dot{V}O_{2max}$ than the targeted 120% of MAS (Dupont et al., 2002), while exercising at 108% of MAS might produce insufficient stimulus to induce the required aerobic adaptions (Myrkos et al., 2022). Highlighting that the TT_{16PLT} is suitable for monitoring changes in aerobic performance, but training prescription should rely on validated MAS tests such as the 30-15 IFT (Grgic et al., 2021), as using TT_{16PLT} derived MAS may lead to inaccurate loads and compromise training effectiveness.

The main limitation of the study is whether the laboratory measure of MAS used can be considered the 'gold standard', as there is a lack of agreement on the most appropriate measurement of MAS in existing literature (Balasekaran et al., 2023). Current literature would suggest there is high variability regarding the speeds and increments used to measure MAS, with some researchers conducting intermittent graded treadmill tests which involve 4-mins of running (starting speed of 2.85 $m{\cdot}s^{-1}$) followed by one minute of rest, increasing by 0.43 $m{\cdot}s^{-1}$ after each rest period (Lacour et al., 1991). Whereas other studies have used a continuous graded treadmill test, determining starting speed at 70% of theoretical HR_{max} (220 minus age in years), with each stage lasting 3-mins and speed increment of 0.42 $m{\cdot}s^{-1}$ (Berthon et al., 1997). In addition, $\dot{V}O_{2max}$ exhibits substantial intergroup variation among individuals, influenced by athletic background and gender, which poses challenges in establishing a universally accepted measure of MAS (Balasekaran et al., 2023). A logistical challenge in the present study was that group sizes during TT_{16PLT} trials were not standardised, ranging from three to twelve participants per session. Previous research suggests that exercising in groups can increase motivation compared to working individually (Nielsen et al., 2014), and anecdotal feedback from participants supported this, with many reporting that running alongside others of similar fitness levels made them protocol feel easier and less fatiguing. Despite these potential influences, the results indicate that group size did not affect performance outcomes. The LoA analysis (Figure 1) showed excellent agreement between repeated trials, with a mean bias close to zero (0.01 $m \cdot s^{-1}$), and the ICC of 0.98 (Table 2) further confirms that variations in group size had no measurable impact on MAS results. Furthermore, allowing participants to complete the TT_{16PLT} in groups offers the advantage of better simulating how the test would be administered in a team sport environment, thereby enhancing ecological validity (Chang et al., 2022).

The present study observed a systematic overestimation of MAS (0.08 $m \cdot s^{-1}$), likely due to the TT_{16PLT} distance (1462.4 m) being too short and therefore requiring a significant contribution from the anaerobic energy system (Thron et al., 2022). Hence, future research should explore the use of longer distances, such as an 18-pitch length TT (1645.2 m) or even a 20-pitch length TT (1828 m), to enhance the predictive validity of MAS measurements. Furthermore, research is needed to investigate the physiological correlates that determine TT_{16PLT} performance; $\dot{V}O_{2max}$, running economy, and anaerobic threshold have all been identified as limiting factors in running performance (Brandon, 1995). Additionally, similar to generic shuttle run tests, an individuals' turning efficiency (Zamparo et al., 2014) and anaerobic speed reserve (Dardouri et al., 2014) are likely to influence TT_{16PLT} performance. The wide LoA 95% CI (Figure 2) suggest a disparity between the physiological determinants affecting performance in both the TT_{16PLT} and the laboratory treadmill test, warranting further research to identify the underlying differences. Another recommended area for future research involves conducting a longitudinal study to examine the consistency in physiological response to training (Burton et al., 2004), time spent at $\dot{V}O_{2max}$, lactate accumulation, and duration to exhaustion between exercise prescribed using laboratory derived MAS and TT_{16PLT} measured MAS. This study would help identify which method of measuring MAS yields a more consistent homogeneous metabolic response to training (Scharhag-Rosenberger et al., 2010).

Conclusion

The present study found no significant difference in MAS between repeated TT_{16PLT} trials ($p > 0.05$), with excellent test-retest reliability demonstrated by an ICC of 0.98 and negligible systematic bias (0.01 $m \cdot s^{-1}$). These findings indicate that TT_{16PLT} provides a consistent and repeatable measure of MAS, making it a practical, time-efficient, and cost-effective option for monitoring changes in aerobic performance over time. However, TT_{16PLT} derived MAS showed an overestimation compared to laboratory derived MAS (mean difference = 0.08 $m \cdot s^{-1}$) and wide LoA (± 0.59 $m \cdot s^{-1}$), suggesting limited validity for precise exercise prescription. Therefore, while the TT_{16PLT} can be confidently used to track performance trends, prescription of training intensities should be based on validated MAS tests, such as the 30-15 IFT, to ensure accuracy and effectiveness.

Conflict of Interest

The authors declare no conflict of interests.

Acknowledgment

The authors would like to thank the hockey players who took part in this study for their commitment and cooperation throughout the data collection process.

References

Altman, D. G., & Bland, J. M. (1983). Measurement in medicine: The analysis of method comparison studies. *Journal of the Royal Statistical Society Series D: The Statistician, 32*(3), 307–317. https://doi.org/10.2307/2987937

Atkinson, G., & Nevill, A. M. (1998). Statistical methods for assessing measurement error (reliability) in variables relevant to sports medicine. *Sports Medicine, 26*, 217–238. https://doi.org/10.2165/00007256-199826040-00002

Aziz, A. R., Chia, M. Y. H., & Teh, K. C. (2005). Measured maximal oxygen uptake in a multi-stage shuttle test and treadmill-run test in trained athletes. *Journal of Sports Medicine and Physical Fitness, 45*(3), 306.

Baker, D. (2011). Recent trends in high-intensity aerobic training for field sports. *Professional Strength & Conditioning, 22*, 3–8.

Baker, D., & Heaney, N. (2015). Review of the literature normative data for maximal aerobic speed for field sport athletes: A brief review. *Journal of Australian Strength and Conditioning, 23*, 60–67.

Balasekaran, G., Loh, M. K., Boey, P., & Ng, Y. C. (2023). Determination, measurement, and validation of maximal aerobic speed. *Scientific Reports, 13*(1), 8006. https://doi.org/10.1038/s41598-023-31904-1

Baquet, G., Berthoin, S., Gerbeaux, M., & Van Praagh, E. (2001). High-intensity aerobic training during a 10-week one-hour physical education cycle: Effects on physical fitness of adolescents aged 11 to 16. *International Journal of Sports Medicine, 22*(04), 295–300. https://doi.org/10.1055/s-2001-14343

Bellenger, C. R., Fuller, J. T., Nelson, M. J., Hartland, M., Buckley, J. D., & Debenedictis, T. A. (2015). Predicting maximal aerobic speed through set distance time-trials. *European Journal of Applied Physiology, 115*, 2593–2598. https://doi.org/10.1007/s00421-015-3233-6

Berthon, P., Fellmann, N., Bedu, M., Beaune, B., Dabonneville, M., Coudert, J., & Chamoux, A. (1997). A 5-min running field test as a measurement of maximal aerobic velocity. *European Journal of Applied Physiology and Occupational Physiology, 75*(3), 233–238. https://doi.org/10.1007/s004210050153

Bland, J. M., & Altman, D. (1986). Statistical methods for assessing agreement between two methods of clinical measurement. *The Lancet, 327*(8476), 307–310. https://doi.org/10.1016/s0140-6736(86)90837-8

Brandon, L. J. (1995). Physiological factors associated with middle distance running performance. *Sports Medicine, 19*, 268–277. https://doi.org/10.2165/00007256-199519040-00004

Brew, D. J., & Kelly, V. (2014). The reliability of the 1.2 km shuttle run test for intermittent sport athletes. *Journal of Australian Strength and Conditioning, 22*(5), 127–131.

Burton, D. A., Stokes, K., & Hall, G. M. (2004). Physiological effects of exercise. *Continuing Education in Anaesthesia, Critical Care & Pain, 4*(6), 185–188. https://doi.org/10.1093/bjaceaccp/mkh050

Chang, M., Büchel, D., Reinecke, K., Lehmann, T., & Baumeister, J. (2022). Ecological validity in exercise neuroscience research: A systematic investigation. *European Journal of Neuroscience, 55*(2), 487–509. https://doi.org/10.1111/ejn.15595

Cohen, J. (2013). *Statistical power analysis for the behavioral sciences.* Academic Press, 25–368.

Dardouri, W., Selmi, M. A., Sassi, R. H., Gharbi, Z., Rebhi, A., Yahmed, M. H., & Moalla, W. (2014). Relationship between repeated sprint performance and both aerobic and anaerobic fitness. *Journal of Human Kinetics, 40*(1), 139–148. https://doi.org/10.2478/hukin-2014-0016

Dupont, G., Akakpo, K., & Berthoin, S. (2004). The effect of in-season, high-intensity interval training in soccer players. *The Journal of Strength & Conditioning Research, 18*(3), 584–589. https://doi.org/10.1519/00124278-200408000-00034

Dupont, G., Blondel, N., Lensel, G., & Berthoin, S. (2002). Critical velocity and time spent at a high level for short intermittent runs at supramaximal velocities. *Canadian Journal of Applied Physiology, 27*(2), 103–115. https://doi.org/10.1139/h02-008

Dupont, G., Defontaine, M., Bosquet, L, Blondel, N., Moalla, W., & Berthoin, S. (2010). Yo-Yo intermittent recovery test versus the Universite de Montréal Track Test: Relation with a high-intensity intermittent exercise. *Journal of Science and Medicine in Sport, 13*(1), 146–150. https://doi.org/10.1016/j.jsams.2008.10.007

Elferink-Gemser, M. T., Visscher, C., Van Duijn, M. A. J., & Lemmink, K. A. P. M. (2006). Development of the interval endurance capacity in elite and sub-elite youth field hockey players. *British Journal of Sports Medicine, 40*(4), 340–345. https://doi.org/10.1136/bjsm.2005.023044

England Hockey Athlete Loading. (2025). https://assets-eu-01.kc-usercontent.com/d66c6a48-e05a-01b8-e0ec-59ee93833239/d5dee843-e0ec-469a-a096-d5e135eb8d6b/Weekly%20Athlete%20Load%20-%20Portrait%20v02.pdf (Accessed 24 August 2025).

England Hockey. (2026). About England Hockey. https://www.englandhockey.co.uk/governance/about-england-hockey

Fédération Internationale de Hockey. (2023). Rules of hockey: Including explanations – effective from January 2022. https://www.fih.hockey/static-assets/pdf/fih-rules-of-hockey-June23-update.pdf.

Galbraith, A. (2021). Physiological assessment of middle- and long-distance runners. In R. C. Blagrove & P. R. Hayes (Eds.), *The science and practice of middle- and long- distance running* (1st ed., pp. 79–96). Routledge.

Grgic, J., Lazinica, B., & Pedisic, Z. (2021). Test–retest reliability of the 30–15 Intermittent Fitness Test: A systematic review. *Journal of Sport and Health Science, 10*(4), 413–418. https://doi.org/10.1016/j.jshs.2020.04.010

Hinton, P. R. (2014). *Statistics explained.* Routledge, London.

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

Hopkins, W. G. (2007). A spreadsheet for deriving a confidence interval, mechanistic inference and clinical inference from a P value. *Sportscience, 11,* 16–21. https://link.gale.com/apps/doc/A189289447/AONE?u=anon~e34de0fc&sid=googleScholar&xid=d9b3330a

Howley, E. T., Bassett, D. R., & Welch, H. G. (1995). Criteria for maximal oxygen uptake: Review and commentary. *Medicine & Science in Sports & Exercise, 27*(9), 1292–1301. https://doi.org/10.1249/00005768-199509000-00009

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/10.1080/026404196367796

Kellmann, M. (2010). Preventing overtraining in athletes in high-intensity sports and stress/recovery monitoring. *Scandinavian Journal of Medicine & Science in Sports, 20*, 95–102. https://doi.org/10.1111/j.1600-0838.2010.01192.x

Kelly, V., & Wood, A. (2013). The correlation between the 30–15 intermittent fitness test and a novel test of running performance. *Journal of Australian Strength and Conditioning, 21*(S1), 91–94.

Kelly, V., Jackson, E., & Wood, A. (2014). Typical scores from the 1.2 km shuttle run test to determine maximal aerobic speed. *Journal of Australian Strength and Conditioning, 22*(5), 183–185.

Koo, T. K., & Li, M. Y. (2016). A guideline of selecting and reporting intraclass correlation coefficients for reliability research. *Journal of Chiropractic Medicine, 15*(2), 155–163. https://doi.org/10.1016/j.jcm.2016.02.012

Kostiukevych, V., Lazarenko, N., Shchepotina, N., Vozniuk, T., Shynkaruk, O., Voronova, V., Kutek, T., Konnov, S., Stasiuk, I., Poseletska, K., & Dobrynskiy, V. (2021). Factor analysis of special qualities of elite field hockey players. *Sport Mont, 19*(S2), 41-47. https://doi.org/10.26773/smj.210908

Krustrup, P., Mohr, M., Amstrup, T., Rysgaard, T., Johansen, J., Steensberg, A., Pedersen, P. K., & Bangsbo, J. (2003). The Yo-Yo intermittent recovery test: Physiological response, reliability, and validity. *Medicine & Science in Sports & Exercise, 35*(4), 697–705. https://doi.org/10.1249/01.mss.0000058441.94520.32

Kusumoto, H., Ta, C., Brown, S. M., & Mulcahey, M. K. (2021). Factors contributing to diurnal variation in athletic performance and methods to reduce within-day performance variation: A systematic review. *The Journal of Strength & Conditioning Research, 35*, S119–S135. https://doi.org/10.1519/jsc.0000000000003758

Lacour, J. R., Padilla-Magunacelaya, S., Chatard, J. C., Arsac, L., & Barthelemy, J. C. (1991). Assessment of running velocity at maximal oxygen uptake. *European Journal of Applied Physiology and Occupational Physiology, 62*(2), 77–82. https://doi.org/10.1007/bf00626760

Leger, L. A., & Lambert, J. (1982). A maximal multistage 20-m shuttle run test to predict O_2 max. *European Journal of Applied Physiology and Occupational Physiology, 49*(1), 1–12. https://doi.org/10.1007/bf00428958

Lombard, W. P., Cai, X., Lambert, M. I., Chen, X., & Mao, L. (2021). Relationships between physiological characteristics and match demands in elite-level male field hockey players. *International Journal of Sports Science & Coaching, 16*(4), 985–993. https://doi.org/10.1177/1747954121998065

Lord, F., Pyne, D. B., Welvaert, M., & Mara, J. K. (2022). Field hockey from the performance analyst's perspective: A systematic review. *International Journal of Sports Science & Coaching, 17*(1), 220–232. https://doi.org/10.1177/17479541211008903

Midgley, A. W., Earle, K., McNaughton, L. R., Siegler, J. C., Clough, P., & Earle, F. (2017). Exercise tolerance during VO_2max testing is a multifactorial psychobiological phenomenon. *Research in Sports Medicine, 25*(4), 480–494. https://doi.org/10.1080/15438627.2017.1365294

Myrkos, A., Smilios, I., Zafeiridis, A., Iliopoulos, S., Kokkinou, E. M., Douda, H., & Tokmakidis, S. P. (2022). Effects of work and recovery duration and their ratio on cardiorespiratory and metabolic responses during aerobic interval exercise. *Journal of Strength and Conditioning Research, 36*(8), 2169–2175. https://doi.org/10.1519/jsc.0000000000003578

Nielsen, G., Wikman, J. M., Jensen, C. J., Schmidt, J. F., Gliemann, L., & Andersen, T. R. (2014). Health promotion: The impact of beliefs of health benefits, social relations and enjoyment on exercise continuation. *Scandinavian Journal of Medicine & Science in Sports, 24*, 66–75. https://doi.org/10.1111/sms.12275

No, M., & Kwak, H. B. (2016). Effects of environmental temperature on physiological responses during submaximal and maximal exercises in soccer players. *Integrative Medicine Research, 5*(3), 216–222. https://doi.org/10.1016/j.imr.2016.06.002

Noblett, H., Hudson, S., Killey, J., & Fish, M. (2023). The physical and physiological match-play locomotor activity profiles of elite domestic male field hockey. *Journal of Sports Science and Medicine, 22*(2), 273–280. https://doi.org/10.52082/jssm.2023.273

O'Donoghue, P. (2013). *Statistics for sport and exercise studies: An introduction.* Routledge, London.

Perrier, E. T., Pavol, M. J., & Hoffman, M. A. (2011). The acute effects of a warm-up including static or dynamic stretching on countermovement jump height, reaction time, and flexibility. *The Journal of Strength & Conditioning Research, 25*(7), 1925–1931. https://doi.org/10.1519/jsc.0b013e3181e73959

Reilly, T., & Borrie, A. (1992). Physiology applied to field hockey. *Sports Medicine, 14*, 10–26. https://doi.org/10.2165/00007256-199214010-00002

Scharhag-Rosenberger, F., Meyer, T., Gäßler, N., Faude, O., & Kindermann, W. (2010). Exercise at given percentages of VO_2max: Heterogeneous metabolic responses between individuals. *Journal of Science and Medicine in Sport, 13*(1), 74–79. https://doi.org/10.1016/j.jsams.2008.12.626

Tabata, I., Nishimura, K., Kouzaki, M., Hirai, Y., Ogita, F., Miyachi, M., & Yamamoto, K. (1996). Effects of moderate-intensity endurance and high-intensity intermittent training on anaerobic capacity and VO_2max. *Medicine and Science in Sports and Exercise, 28*, 1327–1330. https://doi.org/10.1097/00005768-199610000-00018

Tanaka, H., Monahan, K. D., & Seals, D. R. (2001). Age-predicted maximal heart rate revisited. *Journal of the American College of Cardiology, 37*(1), 153–156. https://doi.org/10.1016/s0735-1097(00)01054-8

Thomas, A., Dawson, B., & Goodman, C. (2006). The Yo-Yo test: Reliability and association with a 20-m shuttle run and VO_2max. *International Journal of Sports Physiology and Performance, 1*(2), 137–149. https://doi.org/10.1123/ijspp.1.2.137

Thron, M., Woll, A., Klos, L., Härtel, S., Ruf, L., Kloss, C., & Altmann, S. (2022). Overestimation of maximal aerobic speed by the Université de Montréal track test and a 1500-m time trial in soccer. *Frontiers in Physiology, 13*, 1023257. https://doi.org/10.3389/fphys.2022.1023257

Wong, P. L., Chaouachi, A., Chamari, K., Dellal, A., & Wisloff, U. (2010). Effect of preseason concurrent muscular strength and high-intensity interval training in professional soccer players. *The Journal of Strength & Conditioning Research, 24*(3), 653–660. https://doi.org/10.1519/jsc.0b013e3181aa36a2

Zamparo, P., Zadro, I., Lazzer, S., Beato, M., & Sepulcri, L. (2014). Energetics of shuttle runs: The effects of distance and change of direction. *International Journal of Sports Physiology and Performance, 9*(6), 1033–1039. https://doi.org/10.1123/ijspp.2013-0258