University of Glasgow | College of Medical, Veterinary & Life Sciences

'An investigation of internal and external load-monitoring in soccer goalkeepers.'

MORGAN WRIGHT

BSC (HONS) PHYSIOLOGY & SPORT SCIENCE

XXXXXXXX

Nairn Scobie
Ole Kemi

Submitted in fulfilment of the requirements of the Degree of Master of Sport Science (Research)

School of Life Science, College of Medical, Veterinary & Life Sciences
University of Glasgow

ABSTRACT

Background: Previous investigations into the changes in training load, using internal and external loading metrics, in soccer for outfielders and in small sample sets of GKs have found benefits of using load monitoring strategies to guide periodization of training across the competition season but are limited by sample size, age or lack of GK-specific metrics.
Aims: There were two key themes which prompted this study, with the overarching main aim being to add to the limited body of research currently available on soccer goalkeepers:

1. To investigate changes in internal and external load-monitoring parameters for goalkeepers throughout a four-day match-day lead-in and match day.
2. To determine the effects of changes in Total Distance, Metres per Minute, Dive and Jump metrics on RPE across each day of the four-day match-day lead-in and match day.

The aims of the study were investigated within a Scottish Premiership soccer team and sought to determine additional information about the soccer goalkeeper which could improve the training and match preparation practices currently in place.

Methods: Goalkeepers were monitored using GPS units during training and games and an RPE questionnaire carried out within 30minutes post-training/games, across a four-day match-day lead-in and match day throughout the season. Statistical analysis was carried out on internal and external loading metrics. Subjects were goalkeepers from 1st team, Development squad, U18s and U16s squads of an elite Scottish Premiership soccer team. Ten subjects were used in the study design, training full time across the academy and 1st team (mean age 22.8years ±8.19 years StDev, height 189.5cm ±5.5cm and weight 77kg ±11.4kg).
Results: Results showed RPE was lowest on MD-1, TD was highest on MD-3. $M.Min^{-1}$ was the same on MD-3 and MD. Effect sizes showed key differences between MD-1 and both MD-4 and MD-3 for RPE; between MD-3 and both MD-2 and MD-1 for TD. RPE showed a weak relationship with TD and Total Dives. Overall, differences between each day of the MDLI and MD were marginal when compared to each other.
Conclusions: Results suggest that GK training physical outputs will vary depending on session content throughout the lead-in to a game. On the day before a game, all metrics were lower except Total Jumps, and Time-to-Feet which was slightly quicker. These results added to the existing literature surrounding game demands for the role of a GK, allow the periodization of training to optimize performance capacity. In response to the aims of the investigation, minimal changes to internal and key external loading metrics were found across the MDLI and MD, with MD-3 showing the most significant differences to the subsequent days leading into the MD. In consideration of the second aim, RPE was found to have a weak relationship with total dives and total distance, suggesting that certain external load monitoring metrics could be used to provide information about the GK perceptions of internal load of training and games during the MDLI.
Further research could investigate differences between role requirements for GKs at different levels of involvement on match-days. Additionally, age differences between GKs in 1st teams and development squads could be investigated across more than one club system.

CONTENTS

ABBREVIATION LIST

GK – Goalkeeper
GPS – Global Positioning System
RPE – Rating of Perceived Exertion
TD – Total Distance
HSR – High Speed Running (>5.5m/s unless otherwise specified)
SSG - Small Sided Games
PL - PlayerLoad$_{TM}$
AU – Arbitrary Units
RTF – Return-to-Feet
HR – Heart Rate
IMA - Inertial Movement Analysis (IMA)
Accels - Accelerations (>3m.s^{-2})
Decels – Decelerations (>3m.s^{-2})
HI – High Intensity
SD – Sprint Distance
M.Min^{-1} – Meters per Minute
$\dot{V}O_{2max}$ - Maximal Rate of Oxygen Uptake
G – Game
MD – Match Day
MDLI – Match Day Lead In
MD-1 – The day of the lead-in immediately preceding a match day
MD-2 – The day of the lead-in two days prior to a match day
MD-3 - The day of the lead-in two days prior to a match day
MD-4 - The day of the lead-in two days prior to a match day
MD+1 - The day after a match day
ES – Effect Size
StDev – Standard Deviation

DEFINITIONS

RPE - Rating of Perceived Exertion (RPE) is the numerical score reported by a player on a scale of 1-10, describing how hard they found the session to be physically. The units are arbitrary.
TD – Total Distance - Total Distance in metres covered by the player during the session. TD is measured from the start of the session until the final whistle at the end of the session.
$M.Min^{-1}$ – Metres per Minute - The total distance performed is divided by the duration of session, to provide a measure of intensity given in average metres covered per minute. Metres per minute gives an initial idea of work rate in terms of distance covered per minute, allowing comparison across different session lengths.

Dives - The following metrics relate to the GK-specific movement of diving, whereby both feet leave the ground and the body lands horizontally; either to the left, right or centre. Dives can be analysed by count and intensity and are split into the following metrics:

I. Total Number of Dives – Cumulation of all dives, including all intensities.
II. Total Number of Low Intensity Dives – Number of dives performed at an arbitrary threshold between 3-6AU, predetermined by the host club.
III. Total Number of Medium Intensity Dives - Number of dives performed at an arbitrary threshold between 6-9AU, predetermined by the host club
IV. Total Number of High Intensity Dives - Number of dives performed at an arbitrary threshold 9-12AU, predetermined by the host club
V. Percentage of Total Dives at Low Intensity
VI. Percentage of Total Dives at Medium Intensity
VII. Percentage of Total Dives at High Intensity

Jumps - The following metrics relate to the GK-specific movement of jumping, analysed by count and intensity. Bandings set at 0-20cm for a low intensity jump, 20-40cm for a medium intensity jump and >40cm for a high intensity jump.

I. Total Number of Jumps – Cumulation of all jumps, including all intensities.
II. Total Number of Low Intensity Jumps – Number of jumps performed <20cm high.
III. Total Number of Medium Intensity Jumps - Number of jumps performed 20-40cm high.
IV. Total Number of High Intensity Jumps - Number of jumps performed >40cm high.
V. Percentage of Total Jumps at Low Intensity
VI. Percentage of Total Jumps at Medium Intensity
VII. Percentage of Total Jumps at High Intensity

Return-To-Feet Time refers to the length of time it takes a GK to return to an upright standing position following a dive. It is measured in seconds (s).
PlayerLoad™ - A metric specifically developed by Catapult in order to summate the movements carried out by a player during a session. This will be affected by jumps and dives, as well as horizontal plane movements like running and accelerating/decelerating.
MDLI – The days between match fixtures where preparations take place for the upcoming game day.

LIST OF FIGURES

LIST OF TABLES

ACKNOWLEDGEMENTS

I would like to offer my appreciation to the following people for all their help over the course of my Master's thesis.

Firstly, I would like to thank my supervisor, Nairn Scobie for his relentless support, depth of wisdom and patience each step of the way, both professionally and academically; without him this would not have been possible.

Thanks also to Celtic FC for hosting my Master's internship during the data collection phase of the degree and for the opportunity to be part of the team during the 2018-2019 season.
I would like to offer a huge thank you to my parents, all of my friends, family and colleagues for the time they have invested in me over the last 5 years and each their individual ways in which they have supported me through this journey.

Finally, thanks to God, for everything.

Authors Declaration

"I declare that, except where explicit reference is made to the contribution of others, that this dissertation is the result of my own work and has not been submitted for any other degree at the University of Glasgow or any other institution."

Printed Name: Morgan Christina Wright

Signature: ________________________

INTRODUCTION

1.1 Soccer Overview and Introduction

1.1.1 Overview

Soccer is a dynamic and tactical sport, which has been constantly growing in popularity and developing to become more intense and elicit exciting performances from teams all over the world. Within each team, positional demands vary across the eleven players on the field, as well as additional players brought on as substitute choices for those in the starting XI, with up to 3 substitutions (at time of study, prior to subsequent substitution rule changes, highlighted on 'www.theifab.com', 2022) for each team made during a 90-minute match. The type of players selected for the team often changes depending on the tactical style of play and formation chosen by the manager. However, the soccer goalkeeper (GK) is always present on the field. There is a growing body of research into soccer, from many angles including a plethora of investigation into the physiological demands across the team in training and in games, including a meta-analysis of research by White et al. (2018). However, GKs are often omitted from classical sports science literature, perhaps due in part to the expected differences in the physiological demands placed on the GK (White et al., 2018) due to their unique role during games and training compared to outfielders. However, as practitioners search for marginal gains to improve performance and gain valuable results, there are more studies being published which focus on the goalkeeper, including a case study of a Dutch GK by Malone et al. in 2018. On some occasions where GKs are included in the research into the full team, the investigation will separate the subjects by position, in order to account for the positional differences in both requirements and outputs between outfielders and GKs.

The demands of a GK vary throughout the whole season, in both training and games, from those of a typical outfielder in research investigating training load, Global Positioning System (GPS) outputs and Rating of Perceived Exertion (RPE), due to the differences in training content and match-day requirements. White et al. (2018) compared physical outputs for GKs and outfielders on match days and found that GKs cover around 50% of the total distance covered by the outfielders. The meta-analysis posits that differences in magnitude of physical outputs produced during matches for GKs are affected by the zonal nature of their position; the GK may be fairly stationary for large periods of time during the game and perform dead-ball actions such as goal kicks. A further requirement for a GK is to be able to command the area around the goal as a priority, in order to make decisive actions which impact games, as suggested by Di Salvo et al. (2008), in a study which investigated the activity profile of GKs in games. The study found that GKs perform less physical activity during match play but that the actions they perform are critical for the outcome of the game. In a game, GKs would seem to cover more distance but have less high velocity movements than they would in training sessions (Malone et al., 2018, White et al., 2018). One example of this may be small-sided game training drills, a common part of sessions which involve GKs. They can face a high volume of close-range shots which require extensive physical output, whereas in some competitive fixtures they may experience only a few shots, on average 8-10 shots per game (White et al., 2018) and so maintaining concentration during phases of reduced activity in the defensive third may become increasingly important as an influential factor, contributing to how demanding a game feels to the individual. The systematic review of goalkeeper game

performances, carried out by White et al. in 2018, detailed the technical requirements for the GKs, such as the expectation of around 8-10 shots per game (making an average of 2-10 saves per game, yet provided no analysis of GK-specific physical parameters during training sessions, including dive or jump metrics. This provides a natural area of further investigation into the physical requirements of GKs which forms a focal point of the present study.
GKs have different role requirements to outfielders, and these requirements vary between matches and training, which may impact their perception of exertion on match days and during different training days over the course of the match-day lead-in. Types of drills, session duration and tactics of play can influence how intense the session will be and therefore how exerting the player would perceive it to be. Research into the soccer goalkeeper should aim to add to the limited body of work currently available, in order to help practitioners to train goalkeepers, enabling them to perform at the very peak of their ability.

A low scoring sport, soccer games at every level can be decided quickly in moments, at any time throughout the 90+ minutes of a game. A study by Alberti et al. in 2019 analysed the goal scoring patterns across games in four major European leagues, finding that as time progressed during the game (split into 15minute sections), more goals were likely to be scored. The most goals were scored in the last 15minutes (13.7, 15.1, 16.2, 17,7, 17.2 and 20.2% for 0-15, 15-30, 30-45, 45-60, 60-75, 75-90 minute intervals respectively). These findings suggest that maintaining concentration up until the final whistle is very important for goalkeepers. It was found that 55.1% of goals were scored in the second half, compared to 44.9% in the first half ($p<0.001$). Overall, in the second half especially, fatigue could be a factor which influences goalkeeper performance and thus impacts the number of goals scored and so the final result of the game. The present study has an opportunity to investigate fatigue using performance monitoring techniques to identify how GKs feel after different types of training sessions and games, including an element of tiredness associated with physical exertion.

A qualitative study by Hughes et al. (2012) focused on determining the main skill sets required for elite footballers. The investigation determined that concentration during competition was the most important psychological attribute considered by a group of interviewed performance analysts. It is vital that the GK maintains good focus and performs well, in order to help the team get the result they need. The previously mentioned study by Di Salvo et al. (2008) found that high intensity actions, such as saving, controlling and clearing the ball, which are carried out by GKs can be very decisive in the final result of games, showing how important the GK is to the success of the team. This was highlighted through discussion on how the number of shots and chances increases as the game progresses, supporting the aforementioned study by Alberti et al. (2019) which looked at increased goal-scoring patterns throughout the game timeline, attributed in part to outfielder physical fatigue. This means that GKs must make game-changing actions in these moments to save incoming shots and therefore are crucial for successful team performances.

1.2 Physiological Demands

1.2.1 Positional Game Demands Comparison

On a match day, a GK is governed by some additional rules which only apply to their position. The goalkeepers are the only players permitted to use their hands during games. They may use both hands and feet to save shots and distribute balls during open play or set pieces. Therefore, the way they play football differs to that of outfielders, resulting in different physiological demands as the GK utilizes these additional specifications. A goalkeeper is expected to do a large amount of ball-striking via goal kicks and throwing for distribution, and power movements like explosive dives will be expected to be higher for a GK than outfielders on match days as they defend the area and the goal, as discussed by White et al. (2018). Due to the difference in role requirements of GKs and outfielders, physical outputs in games have been found to vary. A recent study by Abbott, Brickley and Smeeton (2018) investigated GPS data from 44 professional matches in the Premier League and found that across different positions, physical outputs during games will vary for outfielders, with central defenders covering around the least distance with the lowest top speed, wide attackers hitting highest peak velocities and midfielders covering the most distance. Outfielders also sprint approximately every 90secs, with efforts lasting around 2-4s, separated by recovery periods which they can usually choose themselves, in line with tactical demands (Stolen et al., 2005), however, Malone et al. (2017) found GKs to cover only around 17% of the (High Speed Running) HSR distance covered by outfielders. Intensity of activity has been found to influence the standard of football that is played, with players who qualify for international duty performing around 28% ($p<0.05$) more high intensity running (2.43 v 1.90km) and 58% more sprinting (650 v 410m) than professional players involved in football at a lower standard Mohr et al. (2003). The profile for a GK varied compared to outfielders in a meta-analysis literature study by West in 2018. West compared GKs to outfielders, suggesting that a GK will cover between 5.6km and 6km during a game, with 4025m±440m walking, 1223±256m jogging, 221±90m running, 56±34m high speed running and 11±12m sprinting. The lack of high velocity distance covered by GKs which has been investigated in previous literature suggests that it is not a key metric for consideration of GK physiological demands and so is not a primary avenue of investigation for the present study. West also mentions the high intensity actions previously referenced in Section 1.1; saving, controlling and clearing (Di Salvo et al., 2008) during a save. As outfielders do not typically dive to save incoming shots as a GK would, it is hard to compare these type of high intensity actions to what an outfielder might be expected to do, although West acknowledges that these high intensity actions can be decisive in games and so are of high value to the team's performance. Comparing the available literature on GKs and outfielders highlights the importance of the GK specific movements within the role of a GK on match-days and contrasts the different metrics which are more applicable to the physiological game demands of outfielders.

A systematic review carried out by White et al. in 2018 investigated the typical demands of GKs in matches; around 2-10 saves per match, 4-6km total distance covered, 2 short sprints, 8-14 kicks per match into the opponent's half and an extended pre-match warm-up of around 45-60 minutes, consisting of ball-handling and catching. This compares to outfielders who typically cover between 10-12km total distance, position dependant, with a sprint performed approximately every 90s and rarely kick the ball further than 50m. Goalkeepers are also expected to perform as few as 11 high intensity (HI) accelerations and 5 high intensity decelerations compared to outfield players with up to 14 HI accelerations and 24 HI decelerations, during just one 45-minute half. An acceleration or deceleration is typically considered to be high intensity if the rate of change is at least $3ms^{-2}$. The pre-match warm-up for a GK was not detailed in this investigation but Otte et al. (2020) investigated the structure and coaching themes utilised for GKs' match-day warm-ups, finding that GK warm-ups contain basic technical skill work and application of the four key areas of GK-specific work – Skill Acquisition, Defending the Goal, Defending the Area and Distribution. Warm-ups before games include throwing, kicking and catching, as well as shot stopping in a shooting drill at the end of the warm-up with the team's attacking players.

The physiological demands of the game present the opportunity to perform high velocity movements, such as those previously mentioned by Di Salvo et al. (2008) – saving, controlling and clearing, which are not usually carried out by outfielders. In training, there can be opportunities to practice set pieces, such as corners, free kicks and penalties. These are key times within a match setting where a GK must be well prepared. It could be important that the number of high metabolic efforts performed is lower closer to a match (e.g. MD-1) to allow for ample recovery time and preparation, to help reduce the risk of fatigue either before or during a match, as has been found in research by Moreno-Perez et al. (2020). The study by Moreno-Perez et al. in 2020 investigated external load monitoring for soccer GKs during training and matches, finding a tapered periodization whereby total running distance covered and metabolic loading decreased progressively from MD-4 leading into the match-day, however, all days of the lead-in were found to be lower than the match-day. Meanwhile, accels and decels were greater on training days than match days. This suggests that the loading pattern of a GK is not simply to do less movement in total leading into a game, but instead that different external loading metrics will increase or decrease depending on the session content in preparation for match-days.

1.2.2 Training Demands

Typically, GKs will be involved in a combination of team training and position-specific training during a normal training session. This means that the GKs have time to work with an individual coach on aspects of training which are more GK-specific and then are also involved in the same drills as outfielders during other parts of the session, which provides helpful information about the training demands of a GK. In the team training, GKs fit into the drills prescribed for the benefit of the team in general as with all outfielders, often involving some kind of attack or defense on the goal. Time spent in GK-specific training allows the GKs to work on four main areas: Skill Acquisition, Defending the Goal, Defending the Area and Distribution.

A typical training week will incorporate a range of training sessions across each of the four areas mentioned previously, in order to challenge the GK to develop their abilities and prepare for any scenario which they may face in competitive games. In addition to GK-specific training, the goalkeepers will typically facilitate match-simulation drills like small sided games (SSGs), bringing the outfielders a more realistic challenge of aiming their shots into a goal with a GK. One study by Jara et al. (2019) sought to further the understanding of how physical and tactical demands on footballers can be affected by variation in pitch size of SSGs. It was determined that intensity of physical outputs were lower when the pitch size was larger; more jogging and running were performed when the pitch was smallest (32m x 23m) and the mean number of accelerations and decelerations performed was greater in the smallest pitch compared to medium pitches (50m x 35m) or large pitches (62m x 44m), however, the SSGs played on the largest pitch generated just slightly lower numbers of accelerations and decelerations than the smallest pitch size. It was suggested that the differences in physical outputs could be explained by the increased number of defensive and offensive actions covered by a GK when the pitches are smallest, implying that the manipulation of environmental factors, like pitch size or area per player, influences the physical output of GKs. The findings of this study suggest that the size of pitch in training drills affects outfielders, however, despite the GKs involvement being determined as a key contributing factor to intensity of play for outfielders, there was limited investigation into how the GK is impacted by these sessions. This seems like an oversight, considering the important role that the GKs play in both training and games, highlighted by the study itself. It would be beneficial to consider how GKs are impacted by the typical variations of training drill size that occur throughout the match-day lead-in. The present study has an opportunity to build on the previous work by Jara et al. in 2019 and provide an alternative focus on the conditioning aspects of training from various training stimuli on GKs. The findings of Jara et al. can be taken into consideration when preparing training strategies for GKs and help build the foundation for the present investigation to develop more around the impact of periodisation on GKs.

As noted by Jara et al. (2019) in training, physiological movements include walking, jogging, running, side steps/shuffles, single and double leg jumps and dives, with recovery to feet following the dive. These basic movements extend into complex patterns like accelerations towards an oncoming player during an attack from the opposition or a recovery from a dive into distribution from a quick goal kick. These movement patterns can be proactive or reactive, depending on the context of the game (Shamardin and Khorkavyy, 2015), including actions primarily intended to directly or indirectly defend the goal, control operative space within the defensive half of the pitch and to participate in attacks on the opposition goal. This will be portrayed as offensive or defensive movements like moving up the pitch when the team is in possession and play is in the opposition's half, which will increase total distance covered as a proactive tactical move – closing down space to shorten the game and provide an option to receive the ball from a teammate if needed. This proactive movement is not demanded of the GK in every instance but will be used as a tactical decision if it shows potential benefits in the context of that phase of play. Conversely, a reactive situation for a GK might include a 1v1 duel with an oncoming striker from the opposition. The GK must make tactical and technical decisions to avoid conceding a goal to the opponent, resulting in physiological demands such as increased player load from movements carried out vertically, like jumps to reach incoming shots towards the top of the goal.

The physiological demands of a GK throughout the season have previously been investigated in a case study by Malone et al. (2018) on a GK in the top flight of Dutch football. Mean total distance was found to be 22,472m across a typical GK training week with one match, with a Player Load$_{TM}$ (PL) of 1994 (AU) for the week. Average M.Min^{-1} and number of high accelerations and decelerations (>3m.s^{-2}) were also investigated, although GK-specific jump and dive movements were not included in the study. Across the match-day lead-in, TD and PL were found to be lowest on the day before the game and wellness scores were found to be highest on the match day and three days before the game. However, as this was a case study with only one GK, the findings are not as wide-reaching as further research on a greater number of subjects would be, providing an opportunity for the present study to follow a similar vein of research into the match-day lead-in loading pattern but on a larger group of GKs. Using wellness scores and GPS data covers both internal and external loading well, however, no GK-specific movements, such as jumps and dives, were evaluated. More external load-monitoring techniques could be investigated further in the present study to develop a wider view of GK loading leading into a game.

In comparison to the Malone et al. (2018) case study, a study by Owen et al. (2017) found mean TD of GKs was observed to be significantly lower than outfielders during the in-season microcycling phase. Whilst the absolute distance and HSR distance covered by a GK is expected to be significantly lower than that of an outfielder, previous literature has not sought to identify the contribution of GK-specific movements, like jumps and dives, towards the overall physiological demands of a GK's training week. The absence of such investigations within the field of GK research can lead practitioners to fill in the gaps with biases, assumptions and opinions based on anecdotal evidence. Whilst there is absolutely a place for creative choices within prescriptive practice, increasing the available research on GKs specifically can improve the general understanding of the demands of the position, helping practitioners to make better-informed decisions about GK training when leading into matches. This will better prepare the GKs for matches, which benefits the team as a whole, due to their crucial role during competition.

There has always been a responsibility on goalkeepers to stop incoming shots, however, in certain styles of play the goalkeeper can also be considered as an eleventh outfielder. There can be greater importance placed on distribution and commanding the area, as well as traditional shot stopping. The study by McHale, Scarf and Folker (2012) also found that during a match, 21% of total defensive actions are carried out by GKs, which is more than any other position. This shows a huge involvement in the defensive side of the game for GKs, despite physically having some lower outputs than outfielders, including defenders, typically produce. With this information in mind, it is interesting to gain more insights into how goalkeepers train to prepare for such competitions. A study of the performance of goalkeepers during the 2018 World Cup found that incorrect actions were corelated with defeats, and correct actions correlated with a win ($p<0.001$, $r=0.663$) (Mikikis et al., 2021). Understanding more about the role of the GK allows practitioners to prescribe training regimes which prepare the GKs better for competition. The literature surrounding the technical role of the GK provides an excellent backdrop for future studies to identify the physiological requirements needed to meet those technical and tactical role demands.

1.2.3 Goalkeeper-Specific Training

The previously mentioned study by Hughes et al. (2012) used a qualitative design to define key performance indicators from conversations between expert performance analysts which summarised the technical requirements of soccer players, including goalkeepers. A thematic approach identified seven KPIs within 5 classifications - Physiological, Tactical, Technical-

Attacking, Technical-Defending and Psychological. The investigation included ‘positional play, reaction times, calmness’ and ‘positive distribution and communication’ whilst in possession of the ball and to ‘prevent goals, organise defence and be aware’ out of possession of the ball as key roles for a GK. These responsibilities differed from those expected of outfielders. Other performance indicators which were investigated included physiological demands like height, strength, power, agility, coordination, as well as tactical demands such as vision, organisation, and distribution. Technical demands were split into two categories – defensive and attacking. Psychological requirements were also found to include concentration, motivation, attitude and body language. An overarching theme which recurred throughout the study discussed the integration of psychological, technical/tactical and physical skills for GKs training for optimal game preparation. This informs the present study aiming to determine how internal measures of load may be affected by external loading metrics.

The concept of GK-specific training has been explored by numerous research investigations, including the aforementioned skills analysis by Hughes et al. (2012) and an analysis of power output during GK-specific jumping by Vanrenterghem, Lees and Clercq (2008), with research seeking to help GKs preparing for matches. Skill Acquisition is a type of technical performance conditioning, this area of a GK’s training is highly technique based and is very detailed, emphasized in a study by Tee et al. (2018) which states the importance of coaches working with players to develop key skills on the best days of a lead-in in rugby. This is typically carried out early in the match-day lead-in, using previous examples from the most recent game to guide specific areas of focus during the drills (Raab, 2015).

Defending the Goal is the second main area of GK training where GKs work within a very small area on defending their goal through big movement pattens, such as body shape and feet-setting. These movements require confidence and clarity within a tight area, where quick reactions are important for making saves in game-related situations (Szwarc et al., 2019).

Defending the Area incorporates the area around the goal, focusing on forward and lateral changes of direction into the area surrounding the goal itself, with good communication to close down incoming attacks from the opposition (Padulo et al., 2015).

Distribution involves the decision-making process of ball transfer to other areas of the pitch, and is important for the flow of the game, impacting tempo of play and tactical implications for the rest of the team. (Grehaigne et al., 2001; Seaton and Campos, 2011).

These concepts of GK-specific training provide a helpful foundation for understanding more about what kind of movements are performed, how they are executed and how they might affect the overall training load of GKs throughout a competition week. The present study could look to build upon the role requirements of a GK and how that might inform training sessions across the MDLI. Understanding more about the role of a GK during match play allows training to be tailored to the needs of developing and maintaining these skills. The present study could investigate how typical training days might include actions which allow the GK to perform the skills involved in these areas to be successful in matches and training sessions.

1.3 Performance Monitoring

1.3.1 Rationale for Monitoring Performance

There are a wide range of performance monitoring techniques which have been developed since the integration of modern sport science within team sports and soccer in general. Some of the aforementioned studies use a wide range of internal and external performance monitoring methods to help gain more understanding of the demands of the game and how to train best to prepare. Monitoring GKs during training and matches is interesting and important, partly because of the potential to help minimise injury-risks and to ensure adequate stimuli is provided to maintain and develop performance. GKs must compete with each other for the number one place in the starting XI. This can be of great benefit to all GKs in contention for the spot, as it can encourage a high quality of training, however, this could potentially result in differences in training load, particularly in second choice keepers who may be fighting for game time. It would be of critical importance to identify when a GK is training at an intensity greater than they can cope with, to help reduce injury risks. It would also be helpful to understand the physical outputs associated with certain drills, in order to better periodise training schedules. This could be on a micro-scale for weekly programmes, or season-long conditioning plans which could include progressive overload using mesocycles in similar ways to gym programmes for strength. Professional teams endeavour to maximise training benefits with adequate rest time, to reduce the negative effect of unnecessarily missed training days and the financial implications for clubs of paying players who are injured and unable to train or play. The balance of training enough to prepare well for competition but to also rest enough to maintain peak performance throughout a long season of fixtures is a tough battle to win. Djaoui et al. (2017) note the two first steps to successfully manage training for this level of participation; firstly to design the training programme and secondly to monitor physiological changes caused by training and games. Djaoui et al. (2017) found numerous changes to physiological state after training, including blood, urinary and hormonal variations e.g. blood lactate and salivary immunological status. These factors were affected throughout the season and were indicative of markers of training load which are suggested to be related to injury occurrence and to the overall capacity for performance during the competition season.

1.3.2 Periodization

Commonly across a variety of sports, athletes will plan training around competitions using meso-cycling and micro-cycling. These terms refer to the periodization of training throughout competition seasons and can vary in content and duration. Typically, a microcycle is – a four-day, three-day, two-day or one-day match lead-in. The mesocycle is longer, often with key training targets and runs in blocks of around 4-6 weeks in line with international breaks. The aim is typically to load players to their highest capabilities as well as incorporating de-load weeks, whereby the players can rest and recover to be able to maintain a high intensity of training and games. De-load weeks, where total loading covered during the week is minimized, usually fall around the international breaks, however, if GKs are involved in their national team they may have adjusted rest to give some recovery after their international duty is complete.

A study by Chena et al. in 2021 investigated the training loads of professional soccer players throughout a season of microcycling. This study omitted GKs from analysis but shows a pattern of decreased volume and intensity leading into the match, compared to the match being the largest stimulus of the microcycle. It would be interesting to see how GKs loading would compare to this pattern of training load and to determine how this may change throughout the match-day lead-in.

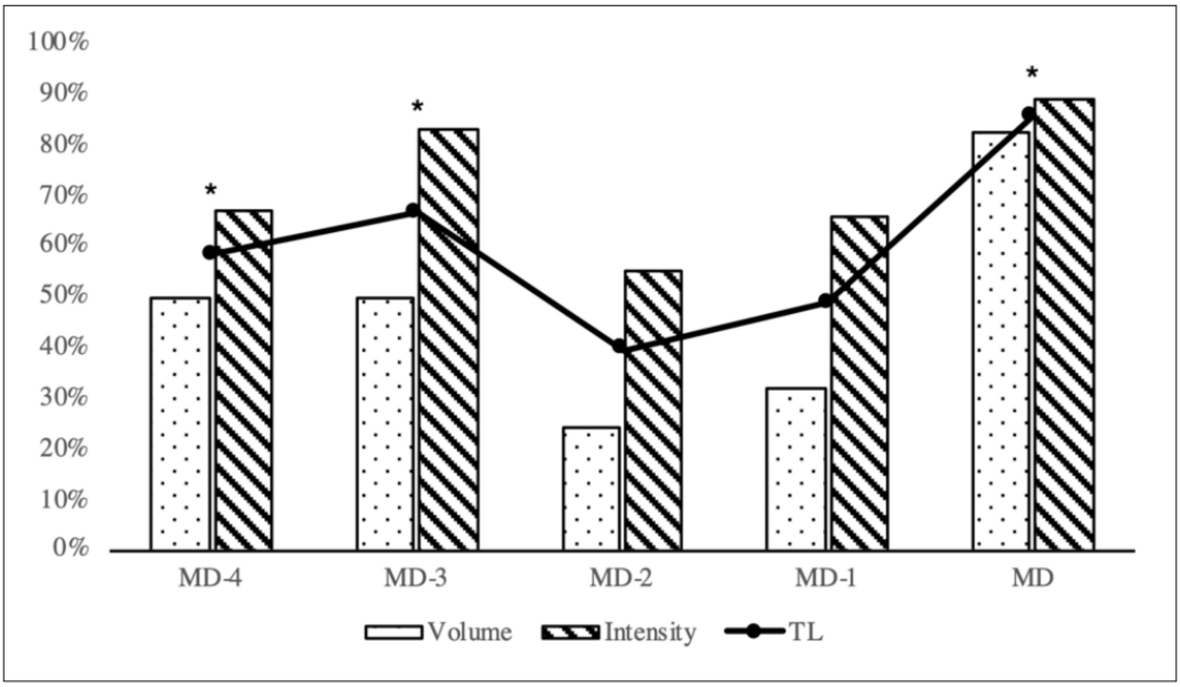

***Figure 1 Outfielder Match-Day Lead-In Loading Pattern**. Presents the loading microcycle of soccer players during a match-day lead-in investigated by Chena et al. (2021).*

An investigation by Malone et al. (2015) which analysed internal and external loading metrics over six 6-week mesocycles across the course of the season found that the training load (including total distance, high speed distance, heart rate and RPE) did not vary across microcycles but that the overall TD was higher during the in-season phase and that positional differences occurred during this phase and during the preseason phase. Figure 2 demonstrates how each shorter microcycle fits into longer mesocycles which make up the whole competition season. Whilst Malone et al. (2015) found total distance and PL to be lowest on the day before the match (MD-1), Chena et al. (2021) found overall training load was lowest two days before the match (MD-2). Although in both cases a taper towards the match day is seen, the differences between MD-2 and MD-1 across studies are an interesting point of potential investigation for the present study.

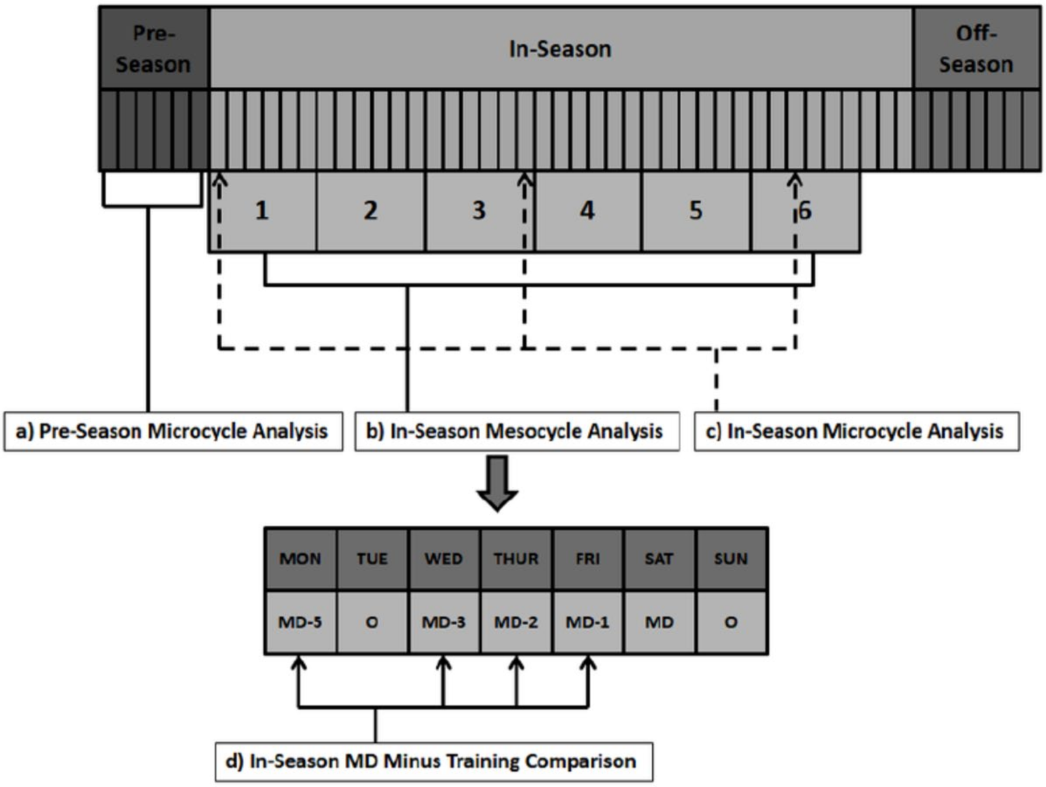

Figure 2 Team Seasonal Loading Pattern. *Presents the seasonal analysis of microcycles and mesocycles for outfielders in the investigation by Malone et al. (2015).*

It is not clear as to why GKs were not included in the Chena et al. study but alongside further investigation by the Malone study, which also did not include GKs, it can be noted that both investigations provide useful information around how the competition season is structured, and both investigations identified similar internal and external metrics to determine the training load of sessions and games. However, both studies excluded GKs from the investigation which makes it difficult to extrapolate the details around training load from outfielders to GKs. It would also be beneficial to understand more around how training requirements of playing GKs compare to those who are not regularly playing but who are still involved in the matchday squad either as warm up GKs or as unused substitutes. The loading systems detailed in Figure 1 and Figure 2 can be used to understand greater seasonal periodisation in professional soccer, in order to elicit performance improvements as well as preparations for competition.

The aforementioned study by Moreno-Perez et al. in 2020 helpfully takes the concepts explored by Chena et al. (2021) and Malone et al. (2015) and applies them in a GK-specific study. Moreno-Perez et al. (2020) present changes in loading metrics, including TD, across the competition week. Similarly to both studies which had focused on outfielders, Moreno-Perez et al. (2020) found that GKs covered most TD during matches. Across all studies, MD-1 can be seen as a low intensity day, in order to aid preparation for the higher intensity match day on the following day. However, some of the explosive loading metrics were found to be higher at the start of the lead-in compared to the MD and can be seen in Figure 3.

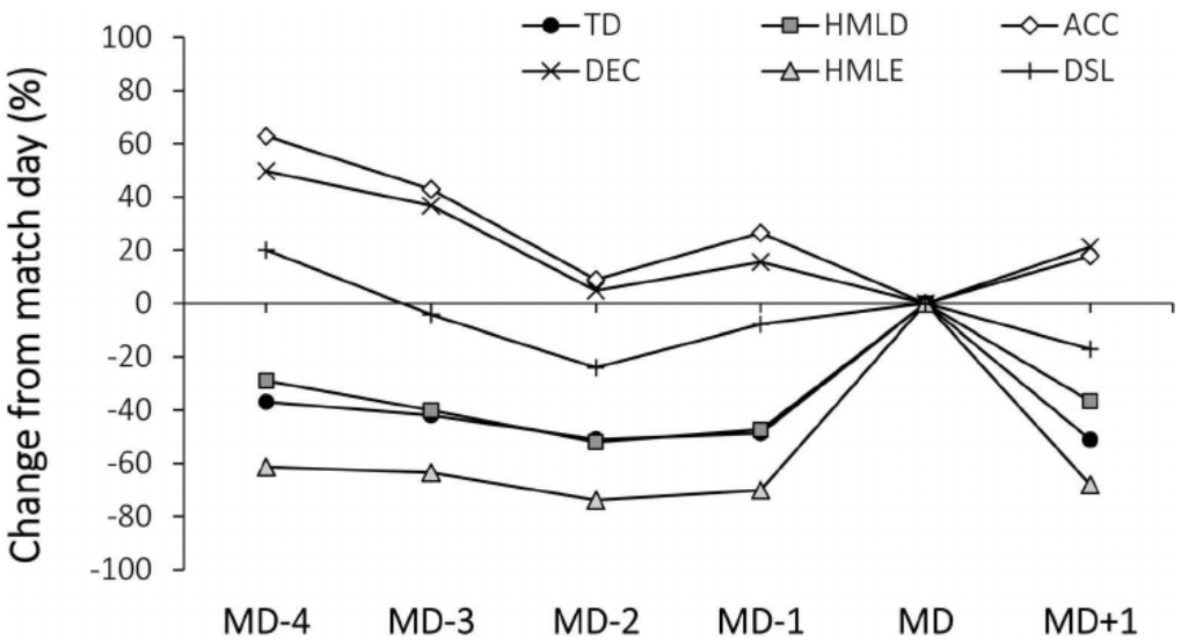

Figure 3 Goalkeeper Match-Day Lead-In Loading Pattern. *Presents the changes from match day (%) to each day of the competition week for external load monitoring metrics (Moreno-Perez et al., 2019).*

A study by Otte et al. (2020) used specialist coaching and skill training periodisation to help develop a youth soccer goalkeeper throughout a competition season. They used coordination training, skill adaptability training and performance training to prepare the GK for competition, identifying movement patterns like stability and variability during the first two phases and then adapting and learning behaviour to integrate team-based training in order to inform performance optimisation in the final stage of the periodisation before competition. The GKs featured in the case study were playing in the Bundesliga youth league in Germany and Figure 4 shows an example of their GK training calendar for the season investigated.

2018	June	July	August	September	October	November	December	
Monday								Monday
Tuesday								Tuesday
Wedensday			1 Training					Wedensday
Thursday			2 Training			1 Saarbrücken (A)		Thursday
Friday	1		3 Training			2 Training		Friday
Saturday	2		4 Waldhof Mannheim (H)	1 Homburg (A)		3 OFF	1 OFF	Saturday
Sunday	3	1 OFF	5 OFF	2 OFF		4 OFF	2 OFF	Sunday
Monday	4	2 Travel TL	6 Training	3 Training	1Training	5 Training	3 Training	Monday
Tuesday	5	3 TL	7 Training	4 Training	2 Training	6 Training	4 Training	Tuesday
Wedensday	6	4 TL	8 Balingen (A)	5 Training	3 OFF	7 OFF	5 Training	Wedensday
Thursday	7	5 TL	9 OFF	6 Training	4 Training	8 Training	6 Training	Thursday
Friday	8	6 TL	10 Training	7 Training	5 Training	9 Training	7 Training	Friday
Saturday	9	7 TEST (Vik. Aschaffenburg (A), 1600)	11 Training	8 Elversberg (H)	6 Training	10 Training	8 Balingen (H)	Saturday
Sunday		8 Travel TL	12 Pirmasens (H)	9 OFF	7 Training	11 Stadtallendorf (H)	9 OFF	Sunday
Monday	11	9 OFF	13 OFF	10 Training	8 Stuttgart II (H)	12 OFF	10 Training	Monday
Tuesday	12	10 Training	14 Training	11 Training	9 Training	13 Training	11 Training	Tuesday
Wedensday	13	11 TEST (Bissingen)	15 Training	12 Training	10 Training	14 Training	12 Training (Test Performance)	Wedensday
Thursday	14	12 Training	16 Training	13 Training	11Training	15 Training	13 Training	Thursday
Friday	15	13 Training	17 Training	14 Walldorf (A)	12 Training	16 Training	14	Friday
Saturday	16	14 TEST (Vik. Koln (A), 1530)	18 Kicker Offenbach (A)	15 OFF	13 Hessen Dreieich (H)	17 Training	15	Saturday
Sunday	17	15 OFF	19 OFF	16 OFF	14 OFF	18 OFF	16	Sunday
Monday	18 Training	16 Training	20 OFF	17 Training	15 Training	19 Training	17	Monday
Tuesday	19 Training	17 Training	21 Training	18 Training	16 Training	20 Training	18	Tuesday
Wedensday	20 Training (Test Performance)	18 Training	22 Training	19 Training	17 Training	21 Training	19	Wedensday
Thursday	21 Training	19 Training	23 Training	20 OFF	18 Training	22 Training	20	Thursday
Friday	22 Training	20 TEST (CfR Pforzheim (A), 1900)	24 Training	21 Training	19 FSV FRA (A)	23 Steinbach (H)	21	Friday
Saturday	23 Training	21 Training	25 Wormatia Worms (H)	22 Training	20 OFF	24 Training	22	Saturday
Sunday	24 OFF	22 OFF	26 OFF	23 Mainz (H)	21 OFF	25 OFF	23	Sunday
Monday	25 Training	23 Training	27 Training	24 OFF	22 Training	26 Training	24	Monday
Tuesday	26 TEST (Arminia Ludwigshafen (H), 1900)	24 Training	28 Training	25 Training	23 Training	27 Training	25	Tuesday
Wedensday	27 Training	25 Training	29 Training	26 Training	24 Training	28 Training	26	Wedensday
Thursday	28 Tests/ Training	26 Training	30 Training	27 Training	25 Training	29 Training	27	Thursday
Friday	29 Tests/ Training - GK COACH MEETING	27 Training	31 Training	28 Ulm (A)	26 Freiburg (H)	30 Mannheim (A)	28	Friday
Saturday	30 OFF	28 Training		29 Training	27 Training		29	Saturday
Sunday		29 TSV Steinbach (A)		30 OFF	28 OFF		30	Sunday
Monday		30 OFF			29 Training		31	Monday
Tuesday		31 Training			30 Training			Tuesday
Wedensday					31 Training			Wedensday

Labels:

Travel/ Training Camp

(TEST) GAME

OFF DAYS

EVENTS

COORDINATION TRAINING

SKILL ADAPTABILITY TRAINING

PERFORMANCE TRAINING

BREAK - Winter/ Sommer (including fitness plan)

Figure 4 Goalkeeper Seasonal Loading Pattern. *Shows the periodisation schedule for GKs at a Bundesliga youth team during 2018-2019 season (Otte et al., 2020).*

Similarly to the present study, this schedule (Figure 4) shows that the number of days leading into a game can vary from week to week. In many cases there were more than 5 days lead in, however, at the start of August there were three consecutive games with only 3 days between the 1st and 2nd games and 3rd and 4th games. This reinforces the importance of understanding what type of training is more physically demanding in the lead-in to games, in order to prevent fatigue impacting competition performances. The use of periodisation may vary across competition levels but is always aimed at optimising performance and preparing athletes to cope with the competition schedule throughout the season. A particular benefit of the study by Otte et al. is the application into youth soccer, which provides useful insight for the present study, compared to the additional body of research investigating different levels of professional soccer.

1.4 Training Load Monitoring

1.4.1 Internal Measures

Internal load monitoring investigates the biochemical (physical and physiological) and biomechanical stress responses to training stimulus, via a variety of monitoring tools, including RPE – which allows the athlete to reflect on their perception of the training session or heart rate parameters like mean sessional heart rate – which is an objective measure of exertion recorded throughout the session (McLaren et al., 2017).

1.4.2 Physiological Markers of Performance Outputs

A systematic review by Hader et al. (2019) found that high speed running (HSR) distance >5.5m/s correlated with biochemical and neuromuscular markers of fatigue, with creatine kinase biomarker activity increasing by 30% and counter movement jump peak power output decreasing by 0.5%, although Total Distance showed no evidence of effects on fatigue-related markers. As GKs are expected to do little to no HSR distance, their highest velocity movements come from high intensity jumps and dives, as well as Return-to-Feet (RtF) time as they recover from dives. Therefore, they may also experience markers of fatigue due to these actions which could impact performance during a match the next day.

Return-to-Feet time is the average time is taken across the session, considering each dive performed by the subject. This time is then divided across the number of dives performed to devise an average length of time taken (in seconds) for the GK to recover from a dive. Bandings are set as shown in Table 1:

Time	Intensity of Recovery
0-0.5s	High
0.5-1s	Medium
>1s	Low

Table 1 Average Time-to-Feet Intensity Bandings. *Presents the categorisation of intensity of recovery from dives, banded by time taken to recover from horizontal dive to upright set position.*

The research by Djaoui et al. (2017) into markers of fatigue addresses the impact of training load on soccer players, finding key internal markers to identify physiological changes after training sessions and on rest days. However, although there were relevant changes to blood, urinary and salivary markers within the investigation, the methods used for sampling these markers were invasive and costly, both in resources and time. Therefore, although the outcome of the research suggested some internal markers were helpful indicators of fatigue, they are not the most practical for use across the MDLI and on match-days within the club environment, due to practicalities for practitioners and athletes, e.g. travelling for competitions meaning that sampling is not always available.

1.4.3 Rating of Perceived Exertion

RPE is a commonly used method of self-assessment among team sports, as it allows for quick, inexpensive and easily collected feedback which can be used alongside session duration to identify an arbitrary 'training load' value. Training load can be monitored longitudinally to aid with competition preparedness. RPE can be reported on a modified Borg scale of 1-10 scale, with 1 being rest and 10 being maximal exertion, although variations exist including differential scales and 15-point scales. Research by Impellizzeri et al., (2004) found RPE could be collected after the session, up to 30 minutes post session, and that RPE was considered a good measure of internal load monitoring for a cohort of youth soccer players but there are conflicting studies which provide guidance on the protocol of collecting scores. A study by Castagna et al. in 2017 collected RPE scores from soccer referees during the FIFA World Cup 2014 Brazil and found that there were no significant differences in scores immediately at the end of the game, after 30mins or after 7hrs post-match, however, the researchers still suggested specific timing of collection would be advised to avoid any issues with recall. Meanwhile in 2022, Abbott and Taber detailed a protocol of collecting RPE scores in young athletes, suggesting a collection time of 15-20mins after the cessation of training or game. These studies, combined with the practicalities of collecting RPE scores after training and games, influence the aim to collect scores as soon as possible after a session or game, but up to 30mins after the end being deemed both possible and acceptable. Type of sport participated in may influence collection period methods but across each study were fairly consistent when considering the practicalities of collecting scores from numerous athletes at the end of their sessions. The ease of use and inexpensiveness of RPE are key reasons for using it as an internal load monitoring tool. In cases where host clubs are already using this as a load-monitoring technique, it also holds a degree of familiarity for athletes which makes it more reliable as a measure of comparison for individual performances. The universal nature of the point and select visual element of the scale shown to athletes also allows for clarity around the question being asked and frequency of use to be appropriate for the busy MDLIs experienced by soccer teams.

The reliability of the modified Borg scale has been tested and promoted by numerous studies, including Impellizerri et al. (2004), Surgenor et al. (2019) and Chen, Fan and Moe (2002), using the scale presented in Figure 11. The subsequent formula used to estimate internal loading (RPE x Session Duration = Training Load) used commonly alongside the collection of such RPE data.

David and Julen (2011) found Player Load$_{TM}$ to correlate with internal load indicators (r=0.331 with mean heart rate percentage and r=0.218 with RPE). however, in any field-based environment it is important to check the ecological validity within the context that the measure is being used. A study by Chen, Fan and Moe (2002) noted that although Borg's RPE scale was a valid tool for measuring exercise intensity, it's correlation with some physiological variables, such as $\dot{V}O_{2max}$ and blood lactate was lower than expected. For goalkeepers, research could provide more clarity on the appropriateness of using RPE measures for assessment of internal load. Surgenor et al. (2019) found significant correlation (r=0.72, p<0.0001) between session-RPE and the Edwards method of calculating training load in a variety of dance activities. The use of session RPE has been investigated alongside heart rate and technical performance of elite youth soccer players during small-sided games (Sanchez-Sanchez, 2017). RPE was found to be higher after game situations with goalkeepers than when no goalkeepers were involved which suggests that GKs play a key role in session intensity, perhaps that outfielders have to work harder when GKs are involved in phases of play. As the GPS technology for GKs has been advancing, the methods of heart rate monitoring technology have increased and so comparing internal load monitoring across both RPE and heart rate may provide further information about GK training and match performances, as with outfielders and in other sports. However, there are some difficulties associated with HR monitoring methods, such as concerns around pain or discomfort for GKs making repeated dives onto chest-mounted or wrist-mounted HR straps. Another challenge of using the available HR monitoring methods for GKs is that during GK-specific movements, such as diving, the straps can become loose and fall off or slip down the abdomen which means that it is not possible to gain accurate HR data. As the safety and comfort of participating GKs, alongside the importance of collecting meaningful and reliable data, is considered of upmost importance in performance monitoring techniques, GKs will not typically wear HR monitoring devices during any football-based sessions. Therefore, finding an alternative means to measure internal loading is imperative for the present study, with the work carried out in the aforementioned investigations by David and Julen (2011) and Sanchez-Sanchez (2017) providing rationale for the use of RPE to assist in the identification of perceived exertion of GKs during different types of sessions across the MDLI.

Previous discussions around the use of heart rate monitoring and as a measure of internal training load suggested a possibility of collecting heart rate (HR) data during training sessions for the potential inclusion within the statistical analysis of this study, however, due to the practicalities of GKs diving during sessions, the heart rate belts available were deemed unfit for purpose. Therefore, heart rate data was not collected for GKs during this investigation. RPE data was used as a means of internal load monitoring, given the depth of previous literature which suggests RPE can be used as a viable method of internal load monitoring instead of average heart rate (Djaoui et al., 2017; Impellizzeri et al., 2004).

There is also a psychological element to RPE which should be considered, as athletes may consider tactical-heavy training sessions to require greater attention to detail and decision-making (Hall, Ekkekakis and Petruzzello, 2005) which may influence the perceived difficulty of the session, without impacting physiological outputs. Hall, Ekkekakis and Petruzzello (2005) found that personality factors can impact RPE, with partial support for the relationship between dispositional and situational psychological factors such as extraversion, behavioural activation and self-efficacy negatively correlating with RPE at lower intensity.

The use of RPE-based training load in soccer was investigated alongside HR to validate the use of RPE to monitor internal load. All correlations showed a strong positive significance between HR-based training load and session RPE of 479 training sessions, with r values ranging from $r = 0.50$ to $r=0.85$ and a p value < 0.01. Training load was calculated by multiplying RPE by session duration in minutes and a number of HR-based training loads were calculated, using Foster et al. (2001) and Edwards' methods (Izzo and Giovannelli, 2017) of assigning relative coefficients to 5 HR zones and summating results. Another method used was Banister's theory of finding the training impulse and the average training session HR and HR at rest. RPE was found to be a useful indicator of internal load and different methods were shown to be helpful for analysing the data provided by the HR monitors (Impellizzeri et al., 2004).

A study by Impellizzeri et al. (2004) investigated the use of RPE-based training load in a group of youth soccer players, in order to understand the role of session RPE as an indicator of global internal load monitoring as part of a training programme. The benefits of using RPE as part of a load monitoring system include the ease of practicality and minimal cost, both of which allow all practitioners a solution to creating a sustainable load monitoring system for soccer training programmes, regardless of budget, time or resources. Another study performed by Rodriquez-Marroyo and Antonan (2015) looked to validate the session RPE using HR monitoring via the Edwards method, similarly to Hall, Ekkekakis and Petruzzello (2005). RPEs were collecting using the Modified Borg Scale and the OMNI scale, which is shown in Figure 11 and has been validated with a range of coefficients from r=0.94 to 0.97 by numerous researchers, including Lagally and Robertson (2006). As a result of the body of work carried out by Rodriquez-Marroyo and Antonan (2015), Hall, Ekkekakis and Petruzzello (2005), Djaoui et al. (2017) and Impellizzeri et al. (2004), session RPE can be considered a valid and reliable method of load monitoring within soccer, and so can be used in the present investigation as a method of internal load monitoring.

1.4.4 External Measures

External load monitoring uses methods outwith the body to determine training intensity, with distance covered being cited as a primary example of an external load monitoring metric by Moreno-Perez et al. (2020). Specifically, this can include GPS parameters like total distance covered and number of accelerations and decelerations. External measures are more objective so reduce bias or individual differences, although some inter unit variability may occur across different GPS units so typically each unit is assigned to an individual player and always used by this individual.

Malone et al. (2018) found PL to be lowest on MD-1 in the aforementioned case study of a Dutch GK across the MDLI. A helpful next step in the research would be to identify how significant the effect of PL is on RPE, alongside other external load-monitoring metrics, to inform the current gaps within the body of literature surrounding GPS technology for GKs.
The present study could build on this through statistical analysis of a mixed model of metrics to identify which metrics might have the strongest relationship with RPE and to what degree this relationship is affected by individual metrics or by the combination of the various physical outputs produced during training for GKs. This could then inform how training sessions could be periodised to best prepare a GK for the match day.

A study by White et al. in 2020 investigated the physical movement profiles of GKs over a week leading into competition. A breakdown of drills suggested that GKs performed most dives during GK-specific training and performed the least dives in SSGs. Investigation of explosive movements, including explosive efforts were found to be higher in GK-specific training than in matches, whilst total distance was highest in games compared to any other type of training. GKs performed less dives in matches than they did in the majority of individual training drills. However, when considering an entire training session to be expected to comprise more than one type of drill, GKs overall load for the measured metrics during training sessions compared to games is harder to quantify, leaving an area of potential further investigation. The present study could follow this vein of research to cumulate training drills into a total training session, across the match-day lead-in, in order to determine how each type of training session compares to a typical match-day load.

A recent study by Casamichana et al. (2024) sought to compare external loading metrics of 3 professional GKs across a microcycle of 4 days leading into competition, as well as the day after the match. It was determined that the total external load decreased in taper towards the match and that MD-4 would be the day of highest external load for the GKs. Although there was no comparison to game days, the MD+1 comparison showed that the external loading metrics - total distance, high speed distance and jumps banded as low, medium, medium-high and high intensity – were higher on MD+1 than on other days of the lead-in, except for MD-4, which could be attributed to 'top up' sessions. These sessions allow GKs to perform additional loading drills to account for time spent inactive on the bench or to increase the load gained through any GK-specific actions which have not been performed during the game. This study provides an interesting take on how training and recovery programmes are periodized for GKs but could benefit from additional GK-specific metrics, such as dives and Return-to-Feet times for GKs performing these role-specific actions during training and games.

1.4.5 GPS Technology

GPS is a system which uses location and time information within a satellite navigation network to track physical outputs of GKs during training and matches. The use of GPS units in professional football is relatively recent (Aughey, 2011), however, GK monitoring remains behind outfielder monitoring, both in variety of available monitoring resources and literature. As a unique position within the sport, GKs are expected to perform less total distance (TD), HSR and accelerations (accels) or decelerations (decels) during matches and training, and usually have a lower aerobic capacity than outfielders whilst performing more high velocity actions such as kicks, jumps and dives (White et al., 2018). The style of play embodied by the team and coaching staff may also alter the requirements of the GK, with discrepancies in TD occurring when comparing a traditional 'shot-stopper' with a 'sweeper-keeper'.

Using GPS devices simplifies the available data into user-friendly parameters, such as 'low intensity jumps,' and can in some cases be customised to suit individual player capabilities via pre-set banding to categorise variables on the analysis software provided by Catapult. This can be useful for the GK specific jump and dive metrics, as bandings can be set for 'low' 'medium' and 'high' bands at different heights, or percentages of personal bests to identify how a player is performing compared to their individualised capability. It would be interesting to investigate the parameters associated with jump and dive bandings, as this is a largely unreported area of GK load monitoring but may have implications on the overall role of what type of sessions during a MDLI would be the most demanding for a GK.

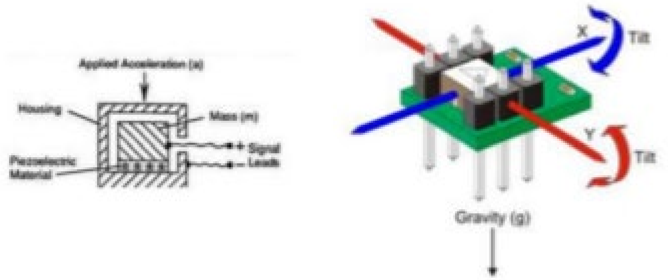

***Figure 5 Triaxial Accelerometer used in Global Positioning System units.** Shows a diagram of the triaxial accelerometer used in the GPS units (Guard, 2017).*

Figure 5 shows how the triaxial accelerometer within the GPS unit uses x, y and z axes alongside gravitational minimisation in order to measure micro movements and direction of movements, referred to as Inertial Movement Analysis (IMA). The GK-specific algorithms, used to determine dive loading metrics, including Return-to-Feet time and dive count, are then added to the GPS units at the end of production. The vertical axis movement is most applicable to a GK jump, perhaps to reach up to catch a cross during a corner. Meanwhile, a short acceleration burst forward to close down space on an incoming attacker would require the x axis along the horizontal axis and a sidestep across the goal might register as a lateral movement along the z axis. A gravity filtering model is applied to compensate for the effects of gravity on the device, preventing inaccuracies between athlete and device movements.
A recent study by Nicollela et al. (2018) investigated the inter- and intra- unit validity of Catapult's S5 GPS units by using the gold standard method for testing accelerations, a shaker table. This found that, compared to the Cartesian formula, PlayerLoad$_{TM}$ was under-reported by around 15%. PlayerLoad$_{TM}$ is a custom GPS metric developed by Catapult to measure external loading on all three axes, horizontal movements like running forwards and backwards, vertical movements like jumping and lateral movements like side steps. The main difference is that a Cartesian formula measures only x and y values as variables of a third value, typically 't' for time. These values act as a parametric wave and in the case of player load, provide forwards and backwards and up and down movements across the horizontal and vertical axis. Therefore, the Cartesian formula would measure only part of the total movement performed by a GK whilst playing or training. Meanwhile, Catapult's 'PlayerLoad$_{TM}$' metric combines triaxial data from vertical, horizontal and lateral movement and uses a unique formula to generate an arbitrary value to represent the total load and 2D load of a player during the session. For outfielders, 2D PlayerLoad$_{TM}$ has been regarded as more accurate for representation of movement impact on the player, as outfield movements occur mainly in horizontal and lateral planes. However, for goalkeepers, many of their movements occur in the vertical plane (for example, jumps). For this reason, PL may be a useful GPS variable for providing information about the physiological cost of movement during both matches and training. If tracked across a MDLI, information around tapering loading could be provided to help prepare GKs better for the demands of match-days. The study carried out by Nicollela et al. (2018) demonstrates how significant the difference can be across methods of calculation, suggesting that particularly a GKs load would be underestimated by a typical Cartesian formula, due to the number of jumps and dives they can make during GK-specific training.

Although GK-specific movements like jumping and diving can be measured individually by the GPS unit, jump height is determined by movement into the air so may not be considered gold standard in accuracy but can provide approximate movement from ground into the air in centimeters. These bandings can be adjusted for individual athletes within the OpenField software but should remain static for the present study, in order to provide an appropriate comparison of actual jump height, rather than relative jump height. Although distinct movements, sometimes, a dive will incorporate a jump. The dive motion can be used in that instance to land the save more safely, although it is not necessarily required to reach the ball or make the save. Some dives require a jump for the GK to attempt a save, due to shot height or GK save technique. In the event of a jump and dive being incorporated into one action, both will be included in count of dive and jump as one individual movement, with further detail provided on the console trace of the Openfield software for each GK.

Figure 6 Goalkeeper Jump Example. *Shows a typical GK jump performed by Aaron Ramsdale (Sport Image, 2020).*

Figure 7 Goalkeeper Dive Example. *Shows a typical GK dive, performed by Aaron Ramsdale (Sport Image, 2020.)*

Figure 8 Combination of Goalkeeper Jump and Dive. *Shows a combination of both jump and dive performed by Aaron Ramsdale to save a shot. This action would be added to both total dive count and total jump count on the metrics collected by the GPS unit. (Sport Image, 2021.*

PlayerLoad$_{TM}$ is measured by the GPS unit and is reported with arbitrary units (AU). With movement forward (anterior-posterior) acceleration, sideways (medial-lateral) acceleration, and vertical acceleration.

However, one limitation of PL is its inability to filter out certain micro-movements like change of directions and high intensity accelerations and decelerations. In order to identify these smaller movements, a combination of accelerometer and gyroscope data with Kalman filtering (an algorithm which uses a large series of sequential measurements, including inaccuracies and noise, to generate estimates), creating a smoothing technique to provide IMA (Catapult White Paper, 2018). The use of both PL and IMA metrics in tandem fills in a more detailed picture of a GK's physical performance, to allow larger scale movements such as Total Distance incorporated into the PL metric as well as tiny micro-movements like variations across dive loading to be identified and analysed, which could be investigated across the MDLI alongside RPE to determine which sessions may have increased load from these movements and could explore their effects on the perceived exertion of GKs preparing for match-days.

1.5 Aims

Throughout the review of literature, key areas of investigation have been common; primarily the use of internal and external load-monitoring metrics to inform soccer goalkeeper training. Some studies focused more on RPE and HR as internal parameters (Rodriquez-Marroyo and Antonan (2015), Hall, Ekkekakis and Petruzzello (2005), Djaoui et al. (2017) and Impellizzeri et al. (2004) providing key investigations in this area), whilst other investigations focused on the use of a variety of external parameters, including TD, $M.Min^{-1}$, PL and GK-specific jump and dive metrics, (Casamichana et al. (2024), White et al. (2020), Moreno-Perez et al. (2020), Malone et al. (2018) providing key studies involving these metrics). However, across all studies, a clear common aim emerged whereby authors sought to enhance the current understanding of GK load monitoring and to provide an evidence-based approach to improving how athletes cope with the demands of the GK role.

Leading on from these findings, there are two key themes to emerge from this study of literature, with the overarching main aim of the present study being to add to the limited body of research available on the role of soccer goalkeepers:

1. To investigate changes in internal and external load-monitoring parameters for goalkeepers throughout a four-day match-day lead-in and match day.
2. To determine the effects of changes in Total Distance, Meters per Minute, Dive and Jump metrics on RPE across each day of the four-day match-day lead-in and match day.

This study will aim to address these themes throughout the investigation and to provide an evidence-based approach to the use of physical load monitoring for soccer goalkeepers.

METHODS AND MATERIALS

Data collection began in July 2018 and continued through until the same time in July 2019. Initial analysis of the data began alongside the collection, due to feedback requirements from coaching staff, however, the bulk of the statistical analysis occurred retrospectively over the course of the timeframe set out for the postgraduate research project and write up. Data was primarily collected by the investigation researcher, however, on occasions when other practitioners were involved in this data collection at times, due to time and practicality constraints, the exact method of collection was provided as a guide to maintain consistent practice. Also, the investigation researcher supervised data collection on multiple occasions prior to any absences, to ensure that the integrity of the data was maintained to the highest possible standard.

Due to the schedule of multiple game weeks, with fixtures often played on different days each week, the MDLI fell on different days of the week throughout the data collection. Therefore, days were considered across the MDLI based on how many days before a game they fall, with the most common consecutive training days typically being 4 in a row. Where there was more fixture congestion, a two-day lead-in was utilised instead.

2.1 Subjects

There were ten GKs at the football club which hosted the postgraduate research project, who were all included in the present study. These GKs range from 16-36 years old (mean 22.8±8.19 years, mean height 189.5±5.5cm and mean weight 77±11.4kg).
GKs played for a variety of teams, including the host club first team (3 GKs), loan club first teams (2 GKs), the development squad (2 GKs), the U18s squad (1 GK) and the U16s squad (2 GKs). Data was only collected from training and matches which GKs played with the host club, although during the 12-month data collection period some GKs went out on loan or played across different squads within the club. All data collected within club training which was viable for analysis were used, including the data collected from players out on loans during training sessions in the club. This is because the loans varied in duration and were a mixture of full-time and part-time, meaning GKs were regularly back within the club wherein this study has been carried out, for training or games. However, data from sessions carried out at loan clubs were not available or included in the analysis.

2.2 Data Collection

2.2.1 External Collection Protocols

GPS units used were 10Hz GPS and 100Hz accelerometer devices - Goalkeeping G5 GPS units from Catapult Sports were labelled, turned on and encrypted to each player prior to session beginning and collected after the session ended. GPS units (shown in Figure 9) were placed into GPS vests by the players or the investigation researcher and were worn in the designated vests, positioned between the scapulae, throughout the session (shown in Figure 10).

. All GPS data was downloaded via OpenField and Catapult software and stored with RPE and TL in Microsoft Excel 2007 databases for analysis. Raw data files were exported after the conclusion of each training session using system-specific software (OpenField, Catapult Sports Ltd., Melbourne, Australia). Individual player files were trimmed to ensure that only information relating to the training time was kept for analysis. In cases whereby GK's were required to undertake any additional training following the conclusion of the main session, this was also accounted.

The GPS units collect data via inertial sensors; accelerometers, GPS, gyroscopes and magnetometers embedded within the device (Catapult White Paper, 2018). The positional data is then calibrated using satellites, with dilution factors showing how many satellites the device is connected to (the more the better). The satellite dilution factor ideally should be at a minimum of 4 for successful connection but ideally will be above 6 (Malone et al., 2017). The horizontal dilution of precision shows the accuracy of GPS as a measure based on geometrical organization of the satellites used for connection and varies from 0-50, with scores below 1AU to be considered ideal, although no set 'gold standards' exist for dilution factors or practical implications for different numbers of satellite connections.

The devices are 100Hz, meaning data is sent between the unit and satellite 100 times per second. Acceleration data is calculated using distance and time, to determine changes in speed, with a smoother algorithm which estimates speed in accordance with anticipated running patterns. The GK algorithms are added to the console, to determine more specific GK movements, like diving and recovering to feet.

Figure 9 Goalkeeper Global Positioning System Unit from Catapult Sports. *Shows a GK GPS unit from Catapult Sports (Catapult White Paper, 2018).*

Figure 10 Goalkeeper Global Positioning System Unit Wearable Positioning. *Shows the GK GPS unit positioned in the custom vest, between the scapulae of a GK from Bournemouth, with GK coach Anthony White during a session. (Catapult White Paper, 2018).*

Collected GPS metrics included the following: Total Distance, M.Min^{-1}, Return-To-Feet Time and PlayerLoad™, as well as dives and jumps under the following categories:

I. Total Number of Dives
II. Total Number of Low Intensity Dives
III. Total Number of Medium Intensity Dives
IV. Total Number of High Intensity Dives
V. Percentage of Total Dives at Low Intensity

VI. Percentage of Total Dives at Medium Intensity
VII. Percentage of Total Dives at High Intensity

I. Total Number of Jumps
II. Total Number of Low Intensity Jumps
III. Total Number of Medium Intensity Jumps
IV. Total Number of High Intensity Jumps
V. Percentage of Total Jumps at Low Intensity
VI. Percentage of Total Jumps at Medium Intensity
VII. Percentage of Total Jumps at High Intensity

2.2.2 Internal Collection Protocols

RPE was collected using a modified Borg scale 1-10 and pen/paper and was then transferred onto the training database on Microsoft Excel 2007. RPE scores were collected verbally using visual Borg scales and recorded within 30 minutes of the session finishing for all training sessions and some games, however, only training data was collected for senior players because the host club had a policy whereby game GPS data was not collected for any professional fixtures, resulting in a very small sample set of game data.

Players were asked 'How hard did you find the session today?' and shown an example modified Borg scale (Figure 11). Their score was noted on the clipboard, privately, in order to prevent other players from potentially seeing their teammates' scores. Inter-practitioner reliability issues were minimized by using a standardized protocol across all data collection, which had already been in place at the host club prior to the beginning of this investigation, allowing great familiarization for both practitioners and subjects. Once scores were transferred onto the Microsoft Excel database, the research investigator was able to perform statistical analysis.

RPE SCALE	RATE OF PERCEIVED EXERTION	EXPLANATION
10	MAX EFFORT ACTIVITY	Feels almost impossible to keep going, unable to talk and completely out of breath. Cannot maintain this level for more than a short time.
9	VERY HARD ACTIVITY	Very difficult to maintain exercise intensity. Can barely breath and speak only a few words.
7-8	VIGOROUS ACTIVITY	Borderline uncomfortable. Short of breath and can just speak in a sentence.
4-6	MODERATE ACTIVITY	Breathing becomes heavier, can hold short conversation. Still somewhat comfortable but becoming noticeably more challenging.
2-3	LIGHT ACTIVITY	Feels possible to maintain for hours. Easy to breath and carry a conversation.
1	RESTFUL/LIGHT ACTIVITY	Hardly any exertion but requires more effort than complete rest.

Figure 11 Modified Borg scale for Rating of Perceived Exertion. *Shows the modified Borg scale used as a visual for GKs providing an RPE after sessions.*

2.2.3 Statistical Analysis

Statistical analysis was carried out using analysis software IBM SPSS Statistics Version 28.0.0.0 (190) and Microsoft Power BI. Tables were created with Microsoft Excel and Power BI.
Data preparation was carried out to by the elimination of incomprehensible outliers and the presentation of sample sizes for metrics on each day of the MDLIs, detailed in Table 2, to identify the data which was available for analysis.

Baseline basic descriptive statistics were then shown for each internal and external metric, with the mean of each parameter, with standard deviations. ANOVA tests were performed with post-hoc Tukey testing showing the 95% confidence intervals, which shown in Table 5. The ANOVAs and Tukey testing were used to determine which physical parameters contribute most to loading on which days throughout the match-day lead-in.

Multiple regression testing was used to identify which external outputs had the greatest influence on RPE. The tables were generated to show standard error, standard deviation and confidence intervals, presenting a pattern of physical loading for each metric throughout the four-day lead-in. The subsequent statistical analysis used a multicollinearity regression model to identify the relationships between parameters. Key metrics have been analysed to highlight potential uses and benefits of monitoring GK loading and also to summarise the potential relationships between internal and external load monitoring factors and the effect the external factors may have on RPE. It was important to perform this test of multicollinearity to ensure closely correlated values would not be simultaneously considered within the same model of regression. The multicollinearity testing allows the model to be cleaned to avoid skewed results; showing which variables will have the biggest impact on how a GK scores their session RPE. Pearson r values of r>0.2 were considered weak, r>0.5 were considered moderate and r>0.7 were considered strong. Metrics with stronger Pearson R values were removed from further analysis in order to avoid interference in the multiple regression analysis, with recognition that a smaller set of key metrics would be investigated as primary metrics.

Basic descriptive statistics were then calculated for the condensed list of metrics which was narrowed down after consideration of the results of multicollinearity testing. These additional descriptive statistics differed from the previous analysis of the established baseline overall combined MDLI and MD means, which can be found in Table 3, because each day of the MDLI and MD was individually analysed to show sample size, means, standard deviations and 95% confidence intervals.

Tables 7, 11, 13, 15, 17 and 19 then details the results of significance testing across each metric on each day of the MDLI and MD. This was carried out to determine which days of the MDLI had the greatest variance.

Effect sizes were then calculated to determine how important the changes leading into the game throughout the match-day lead-in were. Each day of the lead-in, including the match day, was compared to each other day of the lead in, for each metric, to show how important the size of differences were. All days were compared to provide Cohen's d and Glass's delta values where appropriate. When the standard deviations of the compared groups are significantly different, Glass's delta is used. A Levene's test of significance was performed to determine the p value of significant or non-significant differences between standard deviations of each metric on each day of the MDLI and MD, compared with each day of the match-day lead-in and MD. This guided the most appropriate use of either Cohen's d or Glass's delta. Glass's delta is appropriate for comparing samples of significantly different standard deviation and was calculated for effect sizes where necessary. Lin and Aloe (2021) discuss the benefits of using Cohen's d as a primary means of evaluating effect size and note the bands of reference published by Cohen, which allow the size of the calculated effect size to be determined as small, medium or large importance. The bands for this are an effect size =>0.2 as small, =>0.5 as medium and =>0.8 as large. This has been reflected in the analysis provided in Section 3.2.1. Sample sizes for internal and external load monitoring metrics can be found as N values in Table 2. Significance testing was carried out between each day of the MDLI and MD. Where the standard deviations were significantly different, Glass's delta was used. Otherwise, Cohen's d was used. Where appropriate, the relevant effect sizes have been provided in Tables 7, 9, 11, 13, 15, 17 and 19.

Individual models of linear regression were then carried out to identify the relationships between RPE and the key external metrics. For each model, r values of r>0.2 were considered weak relationships, r>0.5 were considered moderate relationships and r>0.7 were considered strong relationships.

To conclude analysis testing, a multiple regression analysis was performed using multivariate linear regression. This was carried out to identify how RPE related to the combination of the key external metrics which had been identified in the multicollinearity testing. These metrics were deemed a key part of the analysis of GK performances and so the relationship between this combination of key metrics and RPE was investigated.

RESULTS

3.1.1 Match-Day Lead-In Session Counts

Data was explored and 5 session entries were removed from further analysis due to incomprehensible errors;

- 4 sessions were removed for showing 0m Total Distance covered, yet registering data in all other metrics.
- 1 session was removed for showing negative Total Distance covered.

Once these outlying data points were removed, further analysis was carried out as described. Table 2 highlights the number of sessions where GPS and RPE were collected. Outliers were removed where GPS had not been worn and on two occasions of incorrect data input on database storage.

Match Day (-)	Measure	N (Sessions)
-4	RPE	46
	GPS	44
-3	RPE	71
	GPS	68
-2	RPE	136
	GPS	109
-1	RPE	173
	GPS	137
G	RPE	72
	GPS	15

***Table 2 Global Positioning System and Rating of Perceived Exertion Data Collection Counts**. Shows the number of sessions analysed for GPS and RPE scores for each day of the match-day lead-in. The 'N' shows the number of sessions included in the investigation and so n is how many sessions analysed were MD-4, MD-3, MD-2, MD-1and MD sessions, i.e. which day of the match-day lead-in each session was on.*

3.1.2 Basic Descriptive Statistics of All Metrics

	Mean	Standard Deviation	Standard Error	95% Confidence Interval
RPE (AU)	3.68	1.34	0.07	(3.54, 3.82)
Total Distance (m)	3031	994	52	(2932, 3132)
$M.Min^{-1}$	36.99	5.94	0.31	(36.39, 37.60)
PL (AU)	309.98	104.12	5.34	(300.05, 320.56)
Total Dives (dives)	34.21	21.10	1.09	(32.10, 36.36)
Low Intensity Dives (dives)	4.05	4.62	0.24	(3.57, 4.53)
Medium Intensity Dives (dives)	14.89	10.60	0.56	(13.80, 15.98)
High Intensity Dives (dives)	15.26	12.31	0.63	(14.10, 16.55)
Total Jumps (jumps)	24.22	13.72	0.71	(22.83, 25.68)
Low Intensity Jumps (jumps)	10.85	8.09	0.41	(10.06, 11.66)
Medium Intensity Jumps (jumps)	10.03	6.96	0.36	(9.33, 10.77)
High Intensity Jumps (jumps)	3.35	4.11	0.22	(2.94, 3.79)
Average Time to Feet (s)	1.62	0.31	0.02	(1.59, 1.65)
% Dives Low (%)	0.13	0.13	0.64%	(0.12, 0.14)
% Dives Medium (%)	0.44	0.17	0.84%	(0.42, 0.45)
% Dives High (%)	0.43	0.20	1.04%	(0.41, 0.45)
% Jumps Low (%)	0.45	0.19	0.99%	(0.43, 0.47)
% Jumps Medium (%)	0.42	0.16	0.82%	(0.4, 0.43)
% Jumps High (%)	0.14	0.16	0.81%	(0.12, 0.15)

Table 3 Results of Descriptive Statistics for All Metrics from Average Across Match-Day Lead-In. *Provides basic descriptive statistics for each internal and external variable measured during the investigation, statistics shown are means, standard deviations, standard errors and 95% confidence intervals.*

In Table 3 the descriptive statistics for all internal and external variables were calculated. Means, standard deviation, standard errors and 95% confidence intervals were calculated based on the variable, irrespective of the day of match-day lead-in.

3.1.3 Multicollinearity

Following on from the general ANOVA analysis of means, correlation coefficients between each variable were calculated, to determine which variables showed the most similarity to each other within the investigation, shown in Table 4.

	RPE	Total Distance	M/Min	PL	Total Dives	Low Intensity Dives	Medium Intensity Dives	High Intensity Dives	Low Intensity Jumps	Medium Intensity Jumps	High Intensity Jumps	Total Jumps	Average Time to Feet	% Dives Low	% Dives Medium	% Dives High	% Jumps Low	% Jumps Medium
RPE																		
Total Distance	0.47																	
M/min	0.25	0.40																
PL	0.50	0.94	0.28															
Total Dives	0.44	0.34	0.08	0.42														
Low Intensity Dives	0.14	-0.04	-0.06	0.03	0.54													
Medium Intensity Dives	0.30	0.14	-0.02	0.22	0.85	0.61												
High Intensity Dives	0.43	0.48	0.18	0.53	0.78	0.03	0.37											
Low Intensity Jumps	0.38	0.37	0.04	0.42	0.31	0.03	0.27	0.29										
Medium Intensity Jumps	0.29	0.36	0.05	0.40	0.34	0.15	0.27	0.29	0.43									
High Intensity Jumps	-0.13	-0.03	-0.02	-0.09	0.17	0.14	0.15	0.12	-0.05	0.22								
Total Jumps	0.33	0.39	0.04	0.43	0.41	0.14	0.34	0.36	0.79	0.83	0.38							
Average Time to Feet	0.05	0.11	-0.12	0.12	0.17	-0.09	0.04	0.28	-0.03	-0.02	-0.08	-0.05						
% Dives Low	-0.17	-0.29	-0.14	-0.28	-0.13	0.47	-0.03	-0.37	-0.13	-0.10	0.03	-0.12	-0.15					
% Dives Medium	-0.12	-0.26	-0.09	-0.24	-0.01	0.14	0.38	-0.39	-0.06	-0.04	0.05	-0.04	-0.18	-0.08				
% Dives High	0.20	0.40	0.16	0.38	0.09	-0.42	-0.29	0.56	0.14	0.10	-0.06	0.11	0.24	-0.58	-0.77			
% Jumps Low	0.21	0.13	0.00	0.16	-0.03	-0.10	-0.04	0.03	0.52	-0.27	-0.52	0.01	0.00	-0.05	-0.07	0.09		
% Jumps Medium	0.02	0.05	0.10	0.07	-0.01	0.03	-0.01	-0.02	-0.33	0.42	-0.19	-0.04	0.05	-0.02	-0.02	0.03	-0.62	
% Jumps High	-0.28	-0.21	-0.10	-0.27	0.04	0.10	0.06	-0.02	-0.31	-0.09	0.84	0.02	-0.05	0.08	0.11	-0.13	-0.62	-0.24

***Table 4 Correlation Coefficients for All Metrics**. Shows correlation coefficients for each variable. Strong correlations are highlighted in red, determined as r=0.7 or below r=-0.7. Weak correlations are highlighted in green, with r values between -0.2 and 0.2.*

Pearson r values of r>0.2 were considered weak, r>0.5 were considered moderate and r>0.7 were considered strong.

Due to particularly strong correlations shown in Table 4, the following variables were removed:

PL and Total Distance had a Pearson r value of 0.94, so PL was removed as this is a multifaceted parameter – accumulating movement on 3 planes and so could be affected in a more complicated way by changes to Total Distance as well as Total Jumps or Total Dives. Due to this, it was decided that Player Load would be the least sensitive measure between PL and Total Distance and so PL was removed.

Total Dives correlated with the following variables: Low Intensity Dives (r=0.54), Medium Intensity Dives (r=0.85) and High Intensity Dives (r=0.78). Total Dives was retained as a variable whilst its components were removed. This is because, for example, if more low intensity dives are performed, the total number of dives will also increase.

Total Jumps had a Pearson r value with the following variables: Low Intensity Jumps (r=0.79), Medium Intensity Jumps (r=0.83) and High Intensity Jumps (r=0.38). Therefore, in similar manner with total dive component metrics, the low, medium and high intensity components of total jumps were also eliminated from the subsequent analysis model.

All percentage metrics were removed due to being calculated derivatives of original variables.

3.1.4 Descriptive Statistics for Match-Day Breakdown of Key Metrics

Basic descriptive statistical testing was performed for each of the shortlisted internal and external metrics; RPE, Total Distance, $M.Min^{-1}$, Total Dives, Total Jumps and Average Time-to-Feet. These results are presented in Table 5.

Match Day (-)		RPE (AU)	Total Distance (m)	M.Min $^{-1}$	Total Dives (dives)	Total Jumps (jumps)	Average Time to Feet (s)
-1	N	173	137	137	137	137	137
	Mean	**3.1**	**2799**	**35**	**34**	**26**	**1.59**
	Std. Deviation	1.14	885	6	18	16	0.31
	95% Confidence Intervals	(2.87,3.25)	(2658, 2937)	(34, 36)	(31, 36)	(24, 29)	(1.53, 1.64)
-2	N	136	109	109	109	109	109
	Mean	**3.46**	**2797**	**38**	**31**	**20**	**1.62**
	Std. Deviation	1.59	973	6	22	15	0.46
	95% Confidence Intervals	(3.48, 4.03)	(2618, 2991)	(37, 39)	(26, 35)	(17, 23)	(1.53, 1.71)
-3	N	71	68	68	68	68	68
	Mean	**4.39**	**3607**	**39**	**37**	**24**	**1.60**
	Std. Deviation	1.12	1024	5	20	15	0.29
	95% Confidence Intervals	(4.20, 4.71)	(3339, 3855)	(37, 40)	(33, 42)	(21, 28)	(1.53, 1.64)
-4	N	46	44	44	44	44	44
	Mean	**4.26**	**3233**	**37**	**40**	**24**	**1.58**
	Std. Deviation	1.22	961	5	29	14	0.31
	95% Confidence Intervals	(3.82, 4.54)	(2964, 3531)	(35, 38)	(32, 48)	(21,29)	(1.48, 1.67)
MD	N	72	15	15	15	15	15
	Mean	**3.61**	**3393**	**39**	**26**	**30**	**1.68**
	Std. Deviation	1.22	919	7	20	14	0.20
	95% Confidence Intervals	(2.93, 4.42)	(2903, 3883)	(36, 42)	(16, 36)	(23, 37)	(1.58, 1.77)

Table 5 Results of Descriptive Statistics for Key Metrics from Each Day of Match-Day Lead-In. *shows basic descriptive statistics for RPE, Total Distance, M.Min^{-1}, Total Dives, Total Jumps and Average Time-to-Feet, including the number of sessions analysed on each day of the match-day lead-in and the mean and standard deviation of each parameter for each day. 95% Confidence Intervals have been provided for all parameters on each day.*

Table 5 has been broken down into its components to provide further information around how each metric has varied across each day of the match-day lead-in. Where confidence intervals do not overlap, there is a significant difference between those days of the MDLI for the specified metric.

3.1.4.i RPE

MD-1 (CI: 2.87,3.25) is significantly lower than MD-2,(CI: 3.48, 4.03) MD-3 (CI: 4.20, 4.71) and MD-4 (CI:3.82, 4.54) . MD-2 (CI: 3.48, 4.03) is significantly lower than MD-3 (CI: 4.20, 4.71). Therefore, on the 2 consecutive days leading into a game, GKs rated session with a lower RPE on average, however, on MD-2, MD-4 and the day of the game, RPE was not significantly different.

3.1.4.ii Total Distance

Total distances covered on MD-2 (CI: 2618, 2991) and MD-1 (CI: 2658, 2937) were significantly lower than MD-3 (CI: 3339, 3855); MD-1 (CI: 2658, 2937) was significantly lower than MD-4 (CI: 2964, 3531). MD-1 and MD-2 can be considered statistically similar to each other and to the game day.

3.1.4.iii M.Min^{-1}

M.Min^{-1} covered was lowest on MD-1 (CI: 34, 36), which was significantly lower than on MD-2 (CI: 37, 39) and MD-3 (CI: 37, 40), however, it was similar to MD-4 and game day, suggesting a peak of distance intensity at the midway point in the lead-in.

3.1.4.iv Total Dives

Total dives performed on each day of the match-day lead-in were not statistically significantly different.

3.1.4.v Total Jumps

Total jumps performed on MD-2 (CI: 17, 23) were significantly lower than on MD-1 (CI: 24, 29), the number of jumps performed on G was higher than all other days of the lead-in.

3.1.4.vi Average Time-to-Feet

The average Time-to-Feet was similar across all days of the match-day lead-in.

3.2 Key Metric Change Across the Match-Day Lead-In

3.2.1 Effect Sizes

The number of sessions where internal and external loading metrics were collected was different across each day of the MDLI, as shown in Table 5, which has implications for the analysis of each metric because comparing metrics with significantly different sample sizes could have greater impact when the sample size is smaller. The standard deviations between means on each day of the MDLI were also different.

Significance testing was carried out between each day of the MDLI and MD. Where the standard deviations were significantly different between specified days of the lead-in, the significance p value has been noted as "(sig)".

Effect sizes were then calculated to determine the changes leading into the game throughout the match-day lead-in. Each day of the match-day lead-in, including the match day, was compared to every other day, to provide either Cohen's d or Glass's delta values where appropriate.

3.2.1.i RPE

The results of the significance testing in Table 6 allow the correct effect size measure to be used.

Day of Match-Day Lead-In	MD-4	MD-3	MD-2	MD-1	MD
MD-4					
MD-3	p=0.51				
MD-2	p=0.04(sig)	p=0.0014(sig)			
MD-1	p=0.53	p=0.88	p<0.0001(sig)		
MD	p=0.98	p=0.48	p=0.01(sig)	p=0.48	

Table 6 Significance Testing for RPE Across Each Day of Match-Day Lead-In and Match Day. *Presents the results of the significance testing between each day of the match-day lead-in and the overall averages for each metric. The"(sig)" denotes a significant difference between standard deviations.*

Effect sizes were then calculated for RPE and are presented in Table 7.

Day of Match-Day Lead-In	-4	-3	-2	-1	MD
-4					
Cohen's d					
Glass's Delta					
-3					
Cohen's d	0.11				
Glass's Delta					
-2					
Cohen's d					
Glass's Delta	0.5	0.59			
-1					
Cohen's d	0.98	1.14			
Glass's Delta			0.23		
MD					
Cohen's d	0.53	0.13		0.43	
Glass's Delta			0.09		

Table 7 Cohen's d and Glass's delta Values for RPE Effect Sizes Across Each Day of Match-Day Lead-In and Match Day. *shows the Cohen's d and Glass's delta values for effect sizes (ES) across the match-day lead-in. Scores =>0.2 have a small effect (highlighted in red), scores =>0.5 have a medium effect (highlighted in amber), scores =>0.75 have a large effect (highlighted in green).*

Table 7 shows a large effect between MD-1 and MD-4 (ES: d=0.98) and between MD-1 and MD-3 (ES: d=1.14). This suggests an important difference between RPE between these two sets of days. MD had a medium effect compared to MD-2 (ES: d=0.53) and MD-2 had a medium effect also with MD-4 (ES: Δ = 0.50), and a medium effect compared to MD-3 (ES: Δ = 0.59) suggesting differences of moderate importance.

3.2.1.ii Total Distance

The results of the significance testing in Table 8 allow the correct effect size measure to be used. All p values were non-significant.

Day of Match-Day Lead-In	MD-4	MD-3	MD-2	MD-1	MD
MD-4					
MD-3	p=0.66				
MD-2	p=0.95	p=0.63			
MD-1	p=0.47	p=0.16	p=0.30		
MD	**p=0.90**	**p=0.68**	**p=0.86**	**p=0.77**	

Table 8 Significance Testing for Total Distance Across Each Day of Match-Day Lead-In and Match Day. *Presents the results of the significance testing between each day of the match-day lead-in and the overall averages for each metric. The"(sig)" denotes a significant difference between standard deviations.*

Effect sizes were then calculated for Total Distance and are presented in Table 9.

Day of Match-Day Lead-In	-4	-3	-2	-1	MD
-4					
Cohen's d					
-3					
Cohen's d	0.38				
-2					
Cohen's d	0.45	0.81			
-1					
Cohen's d	0.47	0.84	0.002		
MD					
Cohen's d	0.17	0.22	0.63	0.66	

Table 9 Cohen's d and Glass's delta Values for Total Distance Effect Sizes Across Each Day of Match-Day Lead-In and Match Day. *Shows the Cohen's d and Glass's delta values for effect sizes across the match-day lead-in. Scores =>0.2 have a small effect (highlighted in red), scores =>0.5 have a medium effect (highlighted in amber), scores =>0.75 have a large effect (highlighted in green).*

Table 9 shows a large effect between MD-2 and MD-3 (ES: d=0.81) and between MD-1 and MD-3 (ES: d=0.84). This suggests an important difference between TD between these two sets of days. MD had a medium effect compared to MD-2 (ES: d=0.63) and MD-1 (ES: d=0.66), suggesting differences of moderate importance.

3.2.1.iii M.Min^{-1}

The results of the significance testing in Table 10 allow the correct effect size measure to be used. All p values were non-significant.

Day of Match-Day Lead-In	MD-4	MD-3	MD-2	MD-1	MD
MD-4					
MD-3	p=0.98				
MD-2	p=0.18	p=0.1			
MD-1	p=0.17	p=0.1	p=0.1		
MD	p=0.09	p=0.07	p=0.37	p=0.36	

Table 10 Significance Testing for M.Min^{-1} Across Each Day of Match-Day Lead-In and Match Day. *Presents the results of the significance testing between each day of the match-day lead-in and the overall averages for each metric. The"(sig)" denotes a significant difference between standard deviations.*

Effect sizes were then calculated for M.Min^{-1} and are presented in Table 11.

Day of Match-Day Lead-In	-4	-3	-2	-1	MD
-4					
Cohen's d					
-3					
Cohen's d	0				
-2					
Cohen's d	0.15	0.18			
-1					
Cohen's d	0.61	0.72	0.5		
MD					
Cohen's d	0.33	0	0.15	0.61	

Table 11 Cohen's d and Glass's delta Values for M.Min^{-1} Effect Sizes Across Each Day of Match-Day Lead-In and Match Day*. Shows the Cohen's d and Glass's delta values for effect sizes across the match-day lead-in. Scores =>0.2 have a small effect (highlighted in red), scores =>0.5 have a medium effect (highlighted in amber). There were no large effect sizes presented.*

Table 11 shows medium effects between MD-1 and all other days of the lead in; MD-4 (ES: d=0.61), MD-3 (ES: d=0.72), MD-2 (ES: d=0.50) and MD (ES: d=0.61). This suggests a moderately important difference for M.Min^{-1} compared to all other days of the lead-in.

3.2.1.iv Total Dives

The results of the significance testing in Table 12 allow the correct effect size measure to be used. Any p values which were significant are denoted in the table.

Day of Match-Day Lead-In	MD-4	MD-3	MD-2	MD-1	MD
MD-4					
MD-3	p=0.01(sig)				
MD-2	p=0.02(sig)	p=0.4			
MD-1	p<0.0001(sig)	p=0.3	p=0.03(sig)		
MD	p=0.13	p=0.93	p=0.72	p=0.52	

Table 12 Significance Testing for Total Dives Across Each Day of Match-Day Lead-In and Match Day. *Presents the results of the significance testing between each day of the match-day lead-in and the overall averages for each metric. The "(sig)" denotes a significant difference between standard deviations.*

Day of Match-Day Lead In	-4	-3	-2	-1	MD
-4					
Cohen's d					
Glass's Delta					
-3					
Cohen's d					
Glass's Delta	0.12				
-2					
Cohen's d		0.29			
Glass's Delta	0.31				
-1					
Cohen's d		0.16			
Glass's Delta	0.21		0.15		
MD					
Cohen's d	0.56	0.55	0.24	0.42	
Glass's Delta					

Table 13 Cohen's d and Glass's delta Values for Total Dives Effect Sizes Across Each Day of Match-Day Lead-In and Match Day. *Shows the Cohen's d and Glass's delta values for effect sizes across the match-day lead-in. Scores =>0.2 have a small effect (highlighted in red), scores =>0.5 have a medium effect (highlighted in amber). There were no large effect sizes presented.*

Effect sizes were then calculated for Total Dives and are presented in Table 13.

Table 13 shows medium effects between MD and MD-4 (ES: d=0.56) and MD-3 (ES: d=0.55). This suggests a moderately important difference for Total Dives on MD compared to these days.

3.2.1.v Total Jumps

The results of the significance testing in Table 14 allow the correct effect size measure to be used. All p values were non-significant.

Day of Match-Day Lead-In	MD-4	MD-3	MD-2	MD-1	MD
MD-4					
MD-3	p=0.82				
MD-2	p=0.82	p=0.99			
MD-1	p=0.6	p=0.56	p=0.49		
MD	p=0.94	p=0.82	p=0.82	p=0.6	

Table 14 Significance Testing for Total Jumps Across Each Day of Match-Day Lead-In and Match Day. *Presents the results of the significance testing between each day of the match-day lead-in and the overall averages for each metric. The"(sig)" denotes a significant difference between standard deviations.*

Effect sizes were then calculated for Total Jumps and are presented in Table 15.

Day of Match-Day Lead-In	-4	-3	-2	-1	MD
-4					
Cohen's d					
-3					
Cohen's d	0				
-2					
Cohen's d	0.28	0.27			
-1					
Cohen's d	0.13	0.13	0.39		
MD					
Cohen's d	0.43	0.41	0.69	0.27	

Table 15 Cohen's d and Glass's delta Values for Total Jumps Effect Sizes Across Each Day of Match-Day Lead-In and Match Day. *Shows the Cohen's d and Glass's delta values for effect sizes across the match-day lead-in. Scores =>0.2 have a small effect (highlighted in red), scores =>0.5 have a medium effect (highlighted in amber). There were no large effect sizes presented.*

Table 15 shows a medium effect between MD and MD-2 (ES: d=0.69). This suggests a moderately important difference for Total Jumps on MD compared to MD-2.

3.2.1.ii Average Time-to-Feet

The results of the significance testing in Table 16 allow the correct effect size measure to be used. Any p values which were significant are denoted in the table.

Day of Match-Day Lead-In	MD-4	MD-3	MD-2	MD-1	MD
MD-4					
MD-3	p=0.62				
MD-2	p=0.0043(sig)	p=0.0001(sig)			
MD-1	p=0.97	p=0.55	p<0.0001(sig)		
MD	p=0.08	p=0.12	p=0.0011(sig)	p=0.07	

Table 16 Significance Testing for Average Time-to-Feet Across Each Day of Match-Day Lead-In and Match Day. *Presents the results of the significance testing between each day of the match-day lead-in and the overall averages for each metric. The"(sig)" denotes a significant difference between standard deviations.*

Effect sizes were then calculated for Average Time-to-Feet and are presented in Table 17.

Day of Match-Day Lead-In	-4	-3	-2	-1	MD
-4					
Cohen's d					
Glass's Delta					
-3					
Cohen's d	0.07				
Glass's Delta					
-2					
Cohen's d					
Glass's Delta	0.09	0.04			
-1					
Cohen's d	0.03	0.03			
Glass's Delta			0.07		
MD					
Cohen's d	0.38	0.32		0.35	
Glass's Delta			0.13		

Table 17 shows minimal effect sizes between all days of the lead in, suggesting no important differences.

Table 17 Cohen's d and Glass's delta Values for Average Time-to-Feet Effect Sizes Across Each Day of Match-Day Lead-In and Match Day*. Shows the Cohen's d and Glass's delta values for effect sizes across the match-day lead-in. Scores =>0.2 have a small effect (highlighted in red), There were no medium or large effect sizes presented.*

3.3 Models of Regression Key Metric Relationships

3.3.1 Key Metrics Models of Regression

Following on from the analysis of multicollinearity and effect sizes, an investigation into the regression model of each parameter with RPE was then performed. This was carried out to determine the significance of each relationship with RPE.

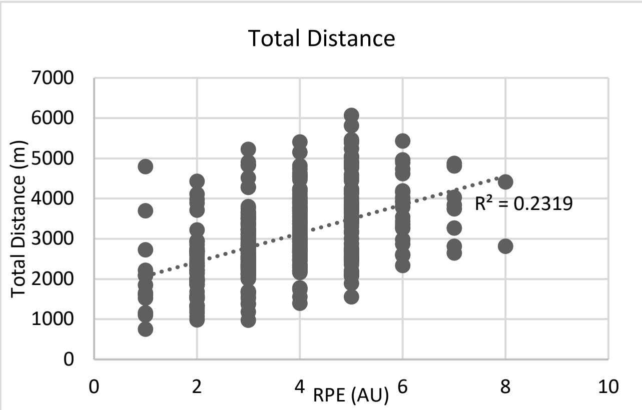

Figure 12 Rating of Perceived Exertion and Total Distance Correlation. *Displays the correlation relationship between RPE and Total Distance.*

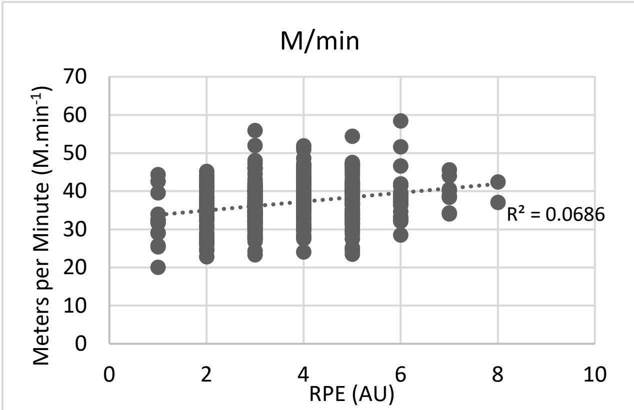

Figure 13 Rating of Perceived Exertion and Meters per Minute Correlation. *Displays the correlation relationship between RPE and M.Min^{-1}.*

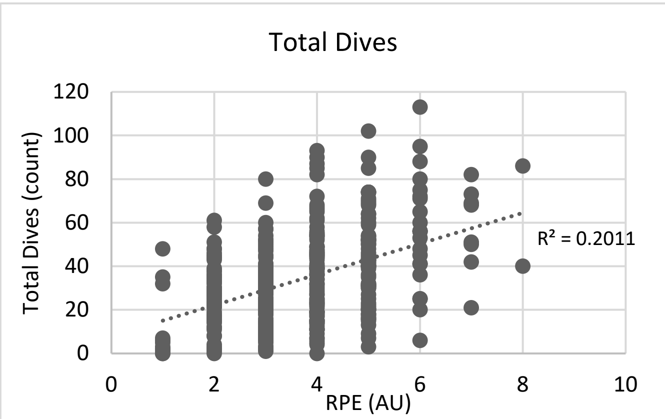

Figure 14 Rating of Perceived Exertion and Total Dives Correlation. *Displays the correlation relationship between RPE and Total Dives.*

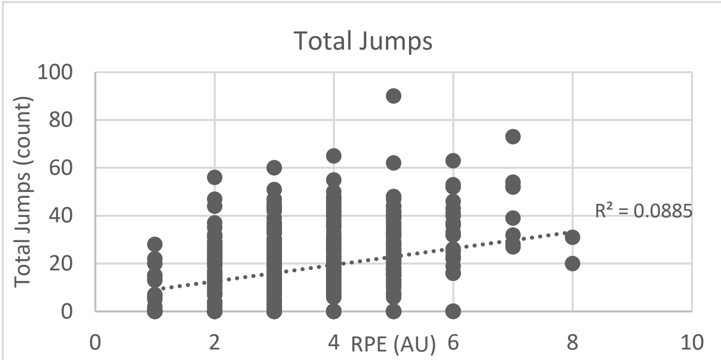

Figure 15 Rating of Perceived Exertion and Total Jumps Correlation. *Displays the correlation relationship between RPE and Total Jumps.*

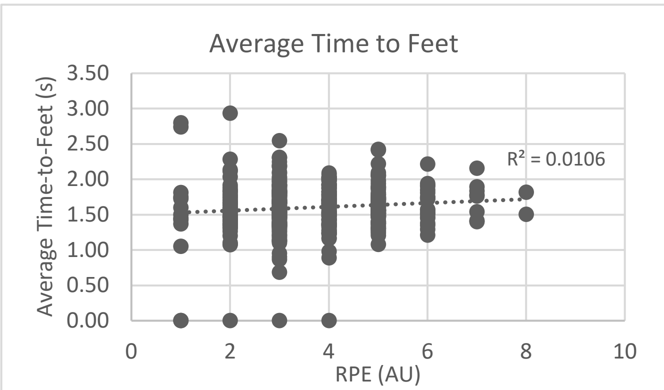

Figure 16 Rating of Perceived Exertion and Average Time-to-Feet Correlation. *Displays the correlation relationship between RPE and Average Time-to-Feet.*

Results from the multiple models of linear regression found all metrics to have an objectively low correlation with RPE, although Total Distance (r^2=0.2319) and Total Dives (r^2=0.2011) had the strongest relationships with RPE from correlation testing. Results showed M.Min^{-1} and RPE to have negligible correlation, R^2=0.0686, Total Jumps to have a negligible correlation with RPE, R^2=0.0885 and Average Time-To-Feet to have a negligible correlation with RPE, R^2=0.0106.

3.3.2 Multiple Regression Model

A multiple regression analysis was performed using multivariate linear regression. When RPE was regressed with a combination of the following variables; Total Distance, M.Min^{-1}, Total Jumps, Total Dives and Average Time-to-Feet, an r^2 value of 0.577 was found. This value suggests a moderate correlation between the mixed model of variables as a whole and RPE.

DISCUSSION

4.1 Aims and Overview

Two main aims were addressed in this investigation:

1. To investigate changes in internal and external load-monitoring parameters for goalkeepers throughout a four-day match-day lead-in and match-day.
2. To determine the effects of changes in Total Distance, Meters per Minute, Dive and Jump metrics on RPE across each day of the four-day match-day lead-in and match-day.

Addressing these themes throughout the investigation and providing an evidence-based approach to the use of physical load management for soccer goalkeepers has highlighted the lack of research into the physical demands of GKs and the real need for more information around how a GK can optimize training strategies to develop the right key skills on the optimal day of the build-up to a game. The heavy fixture scheduling across professional football emphasizes the need for GK load monitoring, as there isn't time to make mistakes around microcycles for competition preparation, due to often very short turnovers between games. In this study, there were more than twice as many MD-2 and MD-1 sessions as there were MD-4 sessions, due to the frequency of two-game weeks and the need for recovery days. This shows how much demand is placed on the GKs to perform consistently at an elite level, throughout a long competition season. Additionally, the demands of GKs to train at the top level to prepare for competition whilst also avoiding injuries is a fine balancing act and requires precise planning to prevent overworking key skills into developing an injury or becoming unavailable through impact injuries. Injuries are a possible common hazard of the position, alongside fatigue-related injuries which were previously investigated in the study on fatigue markers in soccer players by Djaoui et al. (2017). Injuries result in time lost both in games and on the training pitch, and so it is of high importance to condition the GKs to cope with physical demands of the highest level. The present study highlights valuable information around what a typical week looks like for a professional GK, in order to guide the weekly periodisation of training leading into the match-day. This builds on the research discussed previously by Chena et al. (2021), Malone et al. (2015) and Moreno-Perez et al. (2020), which sought to highlight training periodisation models for outfielders and GKs.

The results of this investigation showed that RPE was highest on MD-3 and lowest on MD-1 Total Distance was highest on MD-3 and similarly lowest on MD-2 and MD-1 (both days were very close numerically). $M.Min^{-1}$ was the same on MD-3 and on game days, and numerical very similar on MD-2, suggesting that intensity of movement covered on a MD-3 is most reflective of a typical match-day performance. The highest number of dives performed on any day of the lead-in was on MD-4 compared to a game day with the lowest total dive count. On an average game day, GKs performed the least jumps in total, compared to a MD-2 when the average total jumps were highest.

GKs had very similar Time-to-Feet across each day of the lead in but were slowest on game days compared to the fastest recovery time on MD-4 and MD-1. This may be influenced by MD-1 being primarily a game preparation day in terms of focussing on speed of thought and movement. On MD, mean Total Jumps were highest but mean Total Dives were lowest –

suggesting that even performance outputs of more GK-specific actions differ across the lead-in. The aforementioned investigation by Hughes et al. (2012) similarly found jumping and diving to be key performance outputs of GKs, however, the present investigation has found insight beyond this, to suggest how these skills were utilised across different days of the lead in. All the external variables which were investigated can be closely monitored during goalkeeper training periodization to develop performance and minimize injury risks associated with fatigue.

The results of the statistical analysis show that there are several parameters which may be influenced by the periodisation of the match-day lead-in. RPE and TD presented large effect sizes on MD-3 compared to MD-1, which suggests important differences between outputs for both these metrics on these two days of the lead in. Means and standard deviations were significantly lower on MD-1, which links well to the idea of tapering into a competition, presented by the concept of periodisation in microcycling mentioned previously in this investigation (Otte et al., 2020; Tee et al., 2018). The findings of the present study also link with the study by Malone et al. (2018) which suggested TD would be lower on MD-1 than on MD. Although the present study differs in the specified day of the lead-in on which TD is lowest, there is a point of agreement in that both Malone et al. (2018) and the present study found differences in TD on different days of the competition week. Similarly to Malone et al.'s study which broke sessions down into separate drills, Otte et al. (2020) investigated the unique role of GKs in soccer by exploring GK-specific training sessions and found key skillsets to be considered essential by GK coaches. These skillsets (decision-making, athleticism, mentality and technical skills) have been reflected in the results of the present study; diving and jumping are key aspects of a GKs remit during games and can be seen to be higher on MD-4 and MD-3 of the match-day lead-in, whilst Total Jumps also peaked on MD-1. This is important because it shows that in the present study, GKs have performed significantly more Total Jumps on MD-1 than MD-2 and more Total Jumps were performed on the match days than any other days of the lead-in, partly in order to practice the required skillset laid out by Otte et al. in 2020.

When combined in the individual correlation models with RPE, Total Dives were found to have a slightly positive relationship, showing a possible connection between key GK-specific technical skill practice and the RPE score suggesting how hard a GK is likely to find this type of session. However, the strength of relationship between Total Dives and RPE was weak so should be considered an indication of a potential relationship to be investigated further before implementing into practical settings. Tee et al. (2018) focuses the principles of implementing skill-specific sessions within a holistic approach in rugby union and suggests that every session should incorporate elements of technical and physical skill-training. This point of view relates to the present study because each session carried out across the match-day lead-in contained aspects of each external variable – GKs performed movement on all planes, including jumps and dives every day of training and games. However, as we can see patterns across the lead-ins, with MD-3 typically being a more physically extensive session, some elements of periodisation are retained. This links very well to the research by Tee et al. (2018) in that a holistic approach allows the GKs to practice physical, technical and tactical preparation throughout each session they participate in, although in this case supported by research into a different sport.

However, although the aim may be to incorporate a holistic approach, it may also be important to focus on different elements of training on different days. Similarly to outfielders, for goalkeepers there are different requirements for physical outputs on each day of the match-day lead-in to a game, as discussed by Chena et al. (2021), Malone et al. (2015) and Moreno-Perez et al. (2020). For a typical four-day lead-in, a goalkeeper would be expected to complete certain drills to achieve the required stimulus for these outputs, including various technical, tactical and physical demands. This study found that GKs performed more total distance and a higher meters per minute intensity and rated the session RPE higher on MD-3 than on other days of the match-day lead-in. On match-days, $M.Min^{-1}$ was similar to MD-3 and the number of jumps performed was higher than on any day leading into the game. Average Return-to-Feet time was slowest on game days, which may be due to tactical decisions during key phases of play, such as allowing time for the team to set up to receive goal kicks after a save. Total Distance, $M.Min^{-1}$ and Average Time-to-Feet were low on MD-1 compared to other days of the lead-in. This shows a differentiation between magnitude of physical metrics performed earlier compared with later in the match-day lead-in. This suggests that GKs have some lower physical outputs on last day of the match-day lead-in, which might help reduce the amount of recovery time needed after the session and before the game, as discussed by Chena et al. (2021), Malone et al. (2015) and Moreno-Perez et al. (2020). This was emphasised by the RPE scores being higher on MD-3 compared to MD-1, suggesting that the GKs found the sessions earlier in the lead-in to be more physically demanding than the match-prep session on MD-1, although the variation was minor, effect sizes showed the differences to be most important compared between MD-1 and both MD-4 and MD-3. This again links well to recovery times discussed previously, with regard to tapering into the MD, but also suggests that when the physical outputs of the session are higher, GKs are more likely to rate the session as more physically demanding. It also suggests that MD-1 could be the least physically demanding session of the lead-in, which is ideal when tapering into a game. However, Total Jumps were second highest on MD-1 and highest on the game day. One way in which this research could extend into practical application would be identifying the types of metrics which comprise different drills and creating a library of drills which details the physical outputs likely to be achieved from each type of drill. This could be used to help periodise tougher sessions earlier in the lead-in, to allow more time for recovery, whilst still getting the work into players to improve areas of weakness through specific drills. This would link well to the study by Otte et al. (2020) investigating the microcycling effects of tactical periodisation of GKs by providing very similar outcomes and helpful details around each section of a typical GK session.

Although an increased number of dives were performed on MD-4 compared to other days of the lead in, future research could investigate how RPE changes with even greater increases in the number of jumps and dives performed, especially as the days with highest Total Jumps (MD) and Total Dives (MD-4) were different, perhaps in line with research from Otte et al. in a secondary study published in 2020, which suggests that the GK selected for the starting position on match days will not participate in team shooting drills, because the average number of dives performed during this type of match-day warm up is considered high. It would be beneficial to quantify this and to determine points at which this risk becomes significant enough to modify behaviour.

Previous research by White et al. (2018) may have underestimated the total distance performed by a GK, as previously stated between 4-6km, whilst West (2018) suggested around 5.6-6km, whereas GKs in this study performed much less distance in games (mean of 3.4km), which is more in line with the study performed by White et al. (2018). This may have been partially due to the quality of teams at the host club, who finished in top position in the league, therefore may have been in possession of the ball for longer and thus required less distance from GKs. However, it could be argued that with greater possession, GKs would be able to move further up the pitch and have more freedom to be involved as an 'eleventh outfielder', as previously mentioned. This is a key distinction which could be explored in further research to determine a clearer concept of GK behaviour under different tactical situations.

4.2 Internal Load Monitoring Metrics

Internal load monitoring was carried out via the use of RPE and formed a key part of this investigation. Player perception of exertion was highlighted as an important way to identify how hard sessions were and to understand the impact of heavier and lighter sessions throughout each day of the match-day lead-in. A number of studies (Hughes et al., 2012; Di Salvo et al., 2008; McHale, Scarf and Folker, 2012) discussed at the beginning of this investigation describe key qualities which are important in a GK, the studies focussed not just on physical demands but also described the psychological skills, e.g. concentration during critical moments in games, which are most important for the role. In this study, it was important to use both internal and external metrics, in order to gain insight into the athlete's perception of how hard the sessions were. The reflective nature of RPE scores allow the GK to share the level of exertion they felt was demanded from each session. This is really helpful for understanding how a GK's perception of their fatigue levels vary across sessions, which can help when planning microcycles. GKs may feel more tired after sessions which have higher external load-monitoring metrics; with Total Distance and $M.Min^{-1}$ higher on the same day as the highest RPE score (MD-3) and lowest on the same days as RPE (MD-1 and MD-2). This pattern suggested a link between the total distance and $M.Min^{-1}$ intensity performed by a GK and how fatigued they feel, suggesting that a high magnitude of these external variables could result in an increased RPE score, sharing similarities with the case study by Malone et al. (2018) which found total distance and load derived from RPE to be lowest on MD-1. This may suggest that although not a GK-specific action, the distance covered by a GK and the intensity at which it is covered could be one of the more important indicators of a hard session for a GK. Significance testing also confirmed this in the individual regression model, suggesting that if a GK covers more distance then they may rate the session to be harder than on days when they cover less distance or perform the distance at a lower intensity. This may be because GKs are conditioned for GK-specific movements and are not typically involved in conditioning sessions covering longer distance, due to their specialized positional demands. This may result in a larger incremental impact factor, as performing a less specialized physical output, for which they have not been conditioned for, may be harder. It would be very interesting to explore this area further by investigating GK familiarity with outfielder physical testing, similar to the study by Gil et al. (2007) which investigated maximal and sub-maximal aerobic and anaerobic physiological testing and found significant differences between GKs and other positions, including lowest performances by GKs out of any position for predicted $\dot{V}O_{2max}$ values.

4.2.1 RPE

One aim of the study was to identify if RPE varied across different days, depending on the type of session performed. Although RPE did vary across the lead in, the range of scores was nonetheless quite small, given that RPE is measured on an ordinal 10-point Borg Scale and the range of average means were quite close and had relatively small standard deviations, in comparisons to studies in lab-based settings like Impellizzeri et al. (2004) who used incremental treadmill testing to gauge RPE in the group of young soccer players. In this investigation, GKs seemed to find sessions earlier in the lead-in to be slightly more difficult than those closer to the game. However, because of these differences in RPE being objectively similar, it must be considered that GK training was similarly rated across every day of the MDLI.

Rating of perceived exertion can be influenced by session intensity and duration, suggesting that sessions on days closest to the match would be less intense than those on MD-4 and MD-3, when sessions are set to be harder. This shows that on average, players found MD-3 to be the toughest session, with MD-3 and MD-4 not significantly different from each other. MD-1 was rated by players as significantly less strenuous than all other days in the match-day lead-in. This is ideal for a periodised lead-in, as it allows the GK to recover from harder sessions in time for the competition. Effect sizes shown in Section 3.2 show where differences between days of the MDLI are, with RPE showing important differences between MD-1 and both MD-4 and MD-3. However, there are also moderate differences between RPE on MD-2 and both MD-4 and MD-3, suggesting that while MD-1 and MD-2 are more similar to each other, and MD-4 and MD-3 are more similar, it is important to consider the differences between the 'harder' rated sessions at the start of the lead-in and the 'easier' rated sessions closer to the MD.

The relationship between RPE and player load was supported by findings in David and Julen's study in 2011, with many similarities existing between the present study and the investigation into semi-pro soccer players and workload indicators from GPS units. Although the present study was carried out using a wider range of ages and on GKs instead of outfielders, the results show similarities, particularly in the area of player load. However, due to the correlation testing analysis which showed PL to be very closely correlated with Total Distance, it was removed from this investigation. Therefore, although other studies may use PL as a valid and reliable external load monitoring measure, it could not be justified in the present study beyond initial analysis.

Sanchez-Sanchez et al. (2017) compared RPE after SSGs with and without GKs and among several significant findings, found that outfielder RPE was reported higher in SSGs with GKs than without GKs. This suggests that the perception of exertion can be influenced by the alteration of environmental factors, e.g. the area per player, by including GKs in SSG training situations. This implies an interesting similarity to this study which in turn looks at the GKs perception of their own physical exertion during training when involved in different sessions across the MDLI. The link between conditions of play and GKs perception of fatigue is relevant in the present study, as the types of drills performed on each day of the match-day lead-in vary and RPE is also found to be different throughout the lead-in. Furthermore, with research suggesting an influence of technical and tactical factors on physical outputs, the findings of the present study suggest that RPE will vary across the lead-in as different drills comprise the session. The study by Sanchez-Sanchez et al. (2017) also highlights how the needs of a GK could come second to the requirements of the team in some situations, where GKs may benefit from an alternative training structure to the outfielders, yet may forego this in order to fulfil a wider team drill. Understanding the role of a GK within the team setting and as an individual position is vital for gaining the holistic approach to periodization which helps GKs perform at their best.

The study by Hall, Ekkekakis and Petruzzello (2005) which investigated the psychological aspects of RPE, expecting that athletes may rate tactical-heavy training sessions higher due to mental fatigue, was not supported by the results of this study. This investigation found that on MD-2 and MD-1 (the match preparation sessions become more tactical on the two days leading into the match) RPE was lower than in the sessions earlier in the lead-in, when some physical outputs were higher.

The aforementioned study by Malone et al. (2018) investigated the relationship between self-reported wellness scores and training load, using correlation coefficients to determine the strength of relationship. Similarly, the present study also sought to identify any strength of relationship between key external load monitoring metrics and self-reported RPE scores. Like the study by Malone et al. (2018), the present study found small correlations between TD and RPE and between PL and RPE. While Malone et al. (2018) had found PL to be highest on the same days as TD, and deceleration efforts to have a minimal relationship with wellness scores, the present study found negligible correlations between the two metrics, Total Jumps and Average Time-to-Feet, and RPE. On all fronts the correlations are fairly minimal, with the strongest relationship presented between Total Distance and RPE.

4.3 External Load Monitoring Metrics

The use of periodisation within training originally comes from Vitor Frade’s methodology of tactical periodisation, which has been popular since its introduction around 40 years ago (Bordonau and Villanueva, 2018). Identifying the physical metrics which contribute most to a GK’s sense of fatigue provides very useful information for the planning of microcycles leading into competition. The external parameters identified as most influential in the RPE score given by GKs can then be used within the most appropriate sessions throughout the lead-in, allowing effective tactical periodisation, with a view to combining the different elements of the holistic technical, physical and psychological elements which underpin the tactical work that goes into match-day preparations.

4.3.1 Total Distance and Meters per Minute

Hader et al. (2019) investigated markers of fatigue correlations with HSR distance and found that total distance was not sensitive enough to influence recovery strategies for soccer players. In contrast, this study found higher Total Distance to be linked to highest RPE scores and lower Total Distance to be covered in sessions with lower RPE scores, which suggests a perception of fatigue for GKs in the present study which could be investigated in relation to physiological markers of fatigue as highlighted in the study by Hader et el. (2019) using Creatine Kinase as a predictor of acute and residual fatigue, in order to influence microcycling within tactical periodisation of match-day lead-ins.

Prior discussion of the aforementioned case study by Malone et al. (2018) investigated training day averages for GK external loading metrics, including TD and PL, with results showing them to be lowest on the final day of the match-day lead-in. These results were similar to the present study, which also found TD to be lowest on MD-1. Analysis of the total distances covered by GKs on each day of the lead-in showed that TD on MD-1 and MD-2 was significantly lower than on MD-3. The distances covered on each day are typically lower than for outfielders, due to the varied positional demands, as detailed in the analysis of the study of Owen et al. (2017). An investigation by White et al. (2020) broke down the physiological performance metrics of GKs over a competitive match week, including the match day, and found some differences to the present study, including total distance found to be highest on match days compared to training drills, whilst the present study found the highest TD on MD-3. However, the nature of the study by White et al. (2020) sought to focus on match performance compared to isolated training drills, whilst the present study took into consideration the composition of an entire match-day session content to comprise more than one type of drill. Therefore, although GKs did not cover more TD in any of the individual drills, if the total cumulative distance, or meters per minute, of each day of the lead-in had been quantified, the total distance covered across the whole session for the GK could be expected to be higher than it would be for just one drill in isolation. Also, in the present study, each day of the match-day lead-in and match day was also compared to each other day of the MDLI and MD to find the effect size, suggesting the differences between $M.Min^{-1}$ on MD-1 compared to every other day of the MDLI and MD were moderate. However, these effect size, alongside the others calculated for TD, suggests that although the differences were moderately, they must be considered as objectively negligible, due to the similarity in means across each day.

In the week-long investigation by White et al. (2020) GK-specific movements like jumping and diving were found to be lower during matches than in isolated drill sessions. This would be similar to the present study which also found that GK-specific actions were lower on certain days of the match-day lead-in.

There are several possible roles for a GK on a match-day, with options for GKs to be chosen to start the game, be substitutes on the bench, to be involved in the squad for the warm-up and then any others who are not involved in the match day. Moreno-Perez et al. (2020) explored the differences between loading metrics of GKs who started games and those who did not start, whereas the present study did not separate GKs into these categories, mainly due to restricted access of this information from the host club. Due to the anticipated physical differences expected between the preparation for GKs who are due to start the game compared with non-starter, different lead-ins could be tailored to suit the differing needs of GKs within the squad which should be different for those expected to start games than for those who are not involved in the squad. GKs on the bench could potentially be required from the first minute of play and so should prepare similarly to the starting GK, although physical top-ups could be beneficial after the game if the player does not end up involved in playing in the game, in order to link up loading planning as much as possible. This allows GKs to get as much of a similar match-day stimulus as possible. However, the present study might suggest that although it may be possible to replicate physiological demands of a game in post-match training top ups, there should be an intervention to attempt to provide match stimulus for GKs where possible, due to difficulty in otherwise simulating the variation found in match stimuli, highlighted by the present study finding lowest effect sizes when comparing Total Jumps on across the MDLI and MD. In some cases, the individualisation of GK training and games programme may not be strongly affected by training different GK-specific movements on different days of the MDLI.

Unfortunately, due to the host club's policy, GPS was not used during senior matches, so limited game data was available for analysis within this study. Therefore, the information provided in the meta-analysis by White et al. (2008) is very helpful for understanding the demands of GKs during match-play. GKs involved in the shooting part of the warm-up would be expected to have a higher load than those who were not. Evidence from White et al. (2008) showed typical data produced from GK-specific movements like diving, during match play, suggesting that the GK warm up itself is more physically demanding than a typical game would be. This has implications for loading patterns throughout the rest of the week – the differences between training load for GKs at different levels of MD involvement may call for additional training sessions at other points throughout the MDLI. Results suggest that these top ups may be better placed earlier in the lead-in, e.g. MD-4 or MD-3, where load is expected to be higher and to elicit a higher RPE score than on the days closer to the game where tapering occurs.

Although Total Distance had one of the strongest relationships with RPE, it was still a weak relationship, as shown by the multiple regression analysis. This could suggests that when a GK covers more distance, they will be likely to perceive the session to be harder, however, further investigation should be done to confirm this due to the correlation coefficient still being objectively low. Therefore, on MD-2 and MD-1 it could be beneficial for GKs to cover less distance and lower meters per minute in sessions than earlier in the lead-in, in order to reduce the likelihood of the impact of fatigue effecting match performance.

4.3.2 Dives

The analysis of dive metrics found that the number of dives performed at medium intensity and high intensity were significantly higher than the number of dives performed at low intensity. When the number of dives performed at each intensity were compared as percentages of the total number of dives performed, again the same pattern was found. This suggests that the spread of dive type across the intensity bands during the session was different. This shows that not all dives are the same, there are varying levels of intensity in response to situational factors. From a practical point of view, this information can help coaches understand how players performance outputs will occur in response to different training stimuli. This may be expressed differently on different days of the MDLI.

The average Return-to-Feet metric can be used to average all of the times taken for a GK to recover from a horizontal position back to a standing/set position after each dive during a session. Initially this appeared to be a good option for measuring reactivity and recovery capacity for players. However, with hindsight this exact metric may not be the best for such insights, since not all dives require GKs to recover as quickly as possible. Instead, a more appropriate metric might be ‘median time to feet’, which is also an available metric within the OpenField software. This would statistically ignore the weighted influence of outliers from rested recovery at the end of a set, instead giving a more representative value of the amount of time it took to recover. The method of identifying medians in addition to means has been used previously by Finnoff, Newcomer and Laskowski (2002) to determine the best method for measuring kicking accuracy of soccer players. However, currently software would require a large increase in practitioner labour to derive this information from the data available. There were no significant differences between the mean for average Return-to-Feet time for sessions on each day of the match-day lead-in. Using an alternative metric, such as median return-to-feet time could be more useful in this case, minimising the influence of outliers and providing a different kind of representation for GKs’ ability to recover quickly from dives. However, this metric must be downloaded from the original data export (which was not possible due to restrictions from the host club), as the dive information is determined via the software. In the future, the median Return-to-Feet time could be downloaded as well or instead of the average Return-to-Feet time, to avoid this issue, and such may be considered a limitation of the study. Individual dive analytics must be obtained through visual extraction on the action trace of the console. This is a more time-intensive method which is less practical but can be very beneficial for analysing individual dives. For example, following a match, a successful save and an unsuccessful save can be compared with deeper analytics of pre-dive, impact and post-dive phases. This can be very useful for identifying poor technique and to help correct errors in movement pattern or loading mechanics by comparison to subjectively ‘good’ dives, alongside coaching staff and player feedback also.

4.3.3 Jumps

Although there is very limited literature around GK jumps, a meta-analysis by Hader et al. (2019) identified saving and clearing actions in goalkeeping, which can require a GK to jump, depending on ball height. GKs performed the least jumps during sessions two days before games and most jumps in sessions on the day before games and significantly most on game days. This pattern is quite different to the loading pattern of dives and distance metrics, which all found MD-4 and MD-3 to have higher outputs than the day before a game. This links with the idea that during match preparations on MD-1, session content with outfielders can require GKs to perform more GK-specific actions like jumps, as opposed to covering distance or diving (White et al., 2018).

The recent study by Casamichana et al. (2024) split jump bands up into four separate bands, unlike the present study which used 3 separate bandings and then a Total Jumps band which summated the bandings. The difference between studies is a good example of the flexibility of the GPS analysis system in individualising GK jumps. However, although both studies found jumps to decrease on certain days across the match-day lead-in, the study by Casamichana et al. (2024) determined marginal decreases for jumps at all levels of intensity. Only low intensity jumps had a significant effect size, suggesting an important difference wherein number of jumps on MD+1 was almost double MD-3. Meanwhile, the present study found that only match days had a significant effect size in comparison to the average number of jumps performed on any day of the lead-in, and that the actual number of jumps performed on each day of the lead in were very similar to each other. Although the granularity of detail into jumps in the study performed by Casamichana et al. (2024) was interesting, the lack of match-day data comparison makes it more difficult to understand how these training sessions compare with the competition-specific requirements of a GK. From the information provided in the analysis, it seems as though GKs will perform more jumps during the final session of the match-day lead-in before a game (MD-1), but that the GKs are not likely to perceive a session with more jumps required to be harder than one with less jumps. The present study adds to the body of literature highlighting how training and matches for GKs compare. With that aim a focal point, the present study determines that within this sample group of professional GKs, Total Jumps can be considered a key demand of the GK role during matches and should be form part of the GK preparations throughout the match-day lead-in, with minimal considerations on reducing any fatigue-related effects of training sessions with high volume of Total Jumps, due to the minimal strength of relationship found between Total Jumps and RPE. This is one area where the selected role of the GK can influence their involvement in sessions where a greater total number of jumps may be performed, as GKs not involved in starting or bench performances on the match-day, increased jump exposure may be useful for individual or team development, with minimal risk of increasing the GK's perception of their fatigue.

4.4 Practical Implications

In summary of this investigation, there are a few key points which have been highlighted as important to consider when planning a match-day lead-in for soccer goalkeepers, using internal and external load monitoring.

Firstly, the internal load monitoring metric, RPE, can be used to assign numerical value to how hard a GK rates the session to be. This measure may be a useful tool for assessing player perception of session and to collate with coach targets to determine if the session intensity has matched the desired outcome for each day of the match-day lead-in, however it may lack sensitivity considering the similarity between session ratings across each day of the MDLI. It would be expected that RPE will be similar across each day but that it can still vary with session intensity, which may be better picked up by a more sensitive tool, including other internal markers of fatigue (Djaoui et al., 2017). It was also suggested that there are relationships between RPE and some of the external loading metrics. This would imply that GKs find sessions which have higher total distance, with a greater intensity of meterage per minute to be more physically difficult. Generally, RPE will be lowest during the session on MD-1 and highest during MD-3, which fits well with some of the classical periodisation model of outfielders also.

Secondly, total distance has been presented as an absolute metric and also as distance covered in meters per minute. Distance covered in sessions was around 3km, with MD-1 being the lowest of the match-day lead-in. However, when broken down to $M.Min^{-1}$, MD-3 and MD-2 become statistically similar whilst MD-1 is again the lowest intensity day of the lead-in and is similar to MD-4. This could be linked to drill selection or area of work with less total movement on MD-4 yet it being the second highest RPE-scoring day of the lead-in; the session seems to be intensive, with hard work within smaller areas or less duration.

Thirdly, dive analytics can be useful to identify the number and intensity of dives performed, although these are unlikely to change significantly throughout the match-day lead-in. However, it seems that most dives are carried out at medium intensity or high intensity and a minimal number of dives were performed at low intensity; so the intensity of dive work during sessions was on the higher side, which links back to session RPE and how players felt after performing sessions comprising of large numbers of dives. Dive analytics can show across the lead-in, the intensity of these dives which may be useful for preparing GKs to recover well from more technically or physically challenging dives. This would be especially important when faced with multiple saves during one attacking phase of play or numerous attacking phases in quick succession.

Finally, these internal and external load monitoring metrics have been identified as showing significant or non-significant variation throughout a match-day lead-in but also link to the drill selection which constructs session content, as the GKs will do different types of training on each day of the lead-in, as will outfielders. From this, certain drills could be selected to help GKs reach targets, e.g. avoiding aerobic running drills with the rest of the squad instead to address imbalances across dominant/non-dominant sides during dive recoveries. However, due to total distance being highlighted as a key indicator of RPE, aerobic fitness should not be undervalued and GKs must be properly conditioned for the demands of total distance in both training and games. Examples of drills used during each session could be provided and used in tandem with the findings of the present study, to build a more detailed view of the internal and external loading implications of specific drills and sessions within the match-day lead-in. The GPS metrics available on the GK GPS units can supplement GK load management well and offer an additional means to monitor physical capacity. The use of RPE, total distance, dive and jump metrics particularly can provide information about the session intensity throughout the match-day lead-in. This would suggest that using these metrics in conjunction with periodised loading plans can help identify which sessions are likely to be considered harder or easier for GKs, linking to session content and thus helping to protect or challenge players to maintain and develop physical and technical/tactical ability throughout the competition season, in line with their individual performance plans. The present study provides a blueprint for GK training during a match-day lead-in and identifies useful internal and external loading metrics that contribute to the bigger picture of GK training periodisation. Overall, the literature review, data analysis and subsequent information provided by this study can be used alongside current research to help monitor the internal and external loading of goalkeepers.

4.5 Limitations

Although RPE was found to be significantly lower on sessions immediately prior to competition, the actual score is quite similar on the rating scale, which raises some concern around the sensitivity of the CR10 Modified Borg scale which was used. However, the modified Borg scale remained the primary choice, as players were already familiar with this tool of internal load monitoring within prior post-performance routine, meaning the host club wished to maintain this component of the players' routine throughout the study. However, when considering the variety of session components, perhaps for GKs it would be of more benefit to score their GK-specific training drill, before they are integrated into the session with outfielders, where often they finish with games or shape scenarios which might be considered less physically taxing than GK-specific dive or jump drills. This observation came from conversations with the participants themselves, as well as experience gained from learning about the structure of a GK training session. Also, another reflection from this was that although the GKs were also always asked to provide an RPE score privately, greater care around ensuring player anonymity could be taken to avoid potential issues around social desirability, either amongst their peers and teammates who are competitors for their position or with coaches whose opinions of the player may be influenced negatively by certain RPE scores.

The present study could be developed further to provide detail around the specific drills that each GK participated in throughout the match-day lead-in, to ensure that the sessions actually comprised the physical loading components expected of them. Quantifiable metrics, such as RPE and the external variables investigated in the present study, could provide this detail, alongside subjective coach categorisations which describe key desired outcomes and technical information surrounding drill sizes, number of sets and repetitions or other drill-specific information. To do this effectively, key targets for physical, technical and tactical outcomes should be identified prior to the session, in line with meso- and micro-cycling (Nobari 2021), with reflections after each session to determine whether these drills were successful in achieving targets. With GKs, their sessions can become influenced by the outfielder drills – if the team training requires a GK then they might miss out on more position-specific development or have it rearranged to a less optimal time slot during the lead-in, in order to aid the team as a whole. Knowing what kind of drills elicit certain physical outputs for GKs can help optimal loading throughout the match-day lead-in, allowing development of key GK-specific skills and prioritization of individual needs of the players.

Another area for further investigation could be into individualisation of GK-specific metrics. Similarly, to heart rate and speed bandings which are often individualised to show what percentage of a player's personal best they have achieved, or to the jump bandings used in the study by Casamichana et al. (2024). This could be used with individual jump height personal bests from baseline and in-season testing, to show what percentage of a GK's best ever jump has been performed. This would be a really useful progress tool for developing GKs and could also be used to optimise loading during GK-specific drills, whereby GK movements could be targeted in a similar way to outfielder running drills, e.g. in the Sanchez-Sanchez (2017) study mentioned previously, which used percentage of maximum heart rate achieved to determine training intensity and the subsequent physical loading experienced by players, alongside RPE, in SSGs with or without GKs. This idea could be translated into a GK-specific set of individualised percentage of personal bests, in order to meet the various training demands of different players, who experience different exposure to competition through selection by coaches.

Similarly to studies by Malone et al. (2018) and Otte et al. (2020), but in contrast to systematic reviews by Hader et al. (2019) and White et al. (2018), the present study investigated a sample of players from only one club within one league. In order to produce more widely applicable results, it would be really interesting to extend the sample to include additional GKs at other clubs and in other leagues, especially considering the relatively large range of ages of participants in this investigation. In the study carried out by Otte et al. in 2020, the participants were all development squad players, whilst the case study by Malone et al. (2018) focused on a senior GK in the top Dutch league, Eredivisie; the difference in physical performance outputs of older and younger GKs may vary and so the application of findings could also change depending on the level of competition or style of play across various clubs.

One aspect of the study design which could be improved is the disparity among number of data points on each day of the match day lead in for different metrics. This is mainly due to restrictions in place from the host club; RPE was collected after all sessions but GPS was not allowed to be worn on match days for senior players, due to concerns around player comfort during games and stadium interference with GPS signals. Therefore, the number of MD sessions was very low and not reflective of the full sample of GKs investigated. Also, with such a congested fixture schedule, it was often impossible to have a full four-day match-day lead-in to games, so the number of sessions for MD-4 and MD-3 were lower than that of MD-2 and MD-1. A possible way to combat the issue of disparity among sample sizes of data points could be to conduct longitudinal studies across multiple seasons or with more clubs or teams.

However, in the case of the present study in order to combat the issues around varied number of sessions and data points, both Cohen's d and Glass's delta effect sizes were calculated and presented, to account for differences between lead-in days of similar and different sample sizes. Additionally, because the sample size of subjects was relatively low, a breakdown by GK's age was not carried out. In part, this was preferrable due to the club-wide periodisation strategy which stretched across all professional age groups, meaning that each age group of GKs would have been expected to carry out similar sessions to each other, on the same days of the match-day lead-in. However, as no age breakdown was carried out in the analysis of results in the present study, it is difficult to say if GKs of different ages found the sessions to be equally as hard when rating RPE or had different external outputs in response to the sessions or games in which they took part. Taking age and stage of development into consideration could add another layer of detail to the results produced in the present study.

An additional challenge around data collection presented when the host club prevented any retrospective analysis of the raw data within the console system. This meant that only metrics which had been selected at the beginning of the study were available for analysis.

4.6 Conclusion

When considering the two key aims of this study, it can be concluded that this investigation has indeed added to the limited body of research into soccer goalkeepers and provided additional information around how certain internal and external load monitoring metrics changed across the competition week.

A pattern of internal and external loading has been identified in many of the metrics which were measured and has suggested that on days when GKs perform more of total distance and at a higher intensity of meters per minute, they are more likely to perceive the session to be harder when giving RPE scores. Results also suggested that Average Return-to-Feet Time and Total Jumps did not have any reasonable impact on RPE and were similar across each day of the leading, suggesting that although they are a GK-specific action, the effect on training load is marginal. Additionally, the investigation carried out in the present study sought to highlight the ways in which internal and external load monitoring could be used to benefit practitioners in the tactical periodization of match-day lead-ins for GKs.

Across all the internal and external load monitoring metrics investigated throughout this study, it can be concluded that physical outputs for GKs vary across the four-day match-day lead-in and match, with different parameters presenting higher on certain days of the lead-in compared to others. There also seems to be a relationship between certain external load monitoring metrics and RPE, suggesting that GKs will perceive sessions with increased Total Distance or Total Dives to be harder than those with less. In conclusion, GKs may be fatigued more by certain physical outputs which occur on specific days of the match-day lead-in. This is likely to affect how hard they rate the session, which can influence their preparedness for competition at the end of the match-day lead-in.

REFERENCES

1. Alberti, G., Zangla, D., Siciliano, M., & Piazza, M. (2013). Goal Scoring Patterns in Major European Soccer Leagues. *Sport Sciences for Health, 9*(3), 151–153. doi:10.1007/s11332-013-0154-9
2. Akenhead, R., French, D., Thompson, K. G., & Hayes, P. R. (2014). The acceleration dependent validity and reliability of 10 Hz GPS. *Journal of Science and Medicine in Sport, 17*(5), 562-566.
3. Abbott, W., Brickley, G., Smeeton, N., & Twist, C. (2018). Physical Demands of Playing Position within English Premier League Academy Soccer. *Journal of Human Sport and Exercise, 13*(2). doi:10.14198/jhse.2018.132.04
4. Aughey, R. J. (2011). Applications of GPS technologies to field sports. *International Journal of Sports Physiology and Performance, 6*(3), 295-310.
5. Bahenský, P., Bunc, V., Marko, D., & Malátová, R. (2020). Dynamics of ventilation parameters at different load intensities and the options to influence it by a breathing exercise. *J. Sports Med. Phys. Fit, 60*, 1101-1109.
6. Casamichana, D., Barba, E., Martín-García, A., Ulloa, I., Nakamura, F., & Castellano, J. (2023). Comparison of the external load of professional goalkeepers in different weekly training sessions. *Biology of Sport, 41*(1), 67-82.
7. Castagna, C., Bizzini, M., Póvoas, S. C. A., & D'Ottavio, S. (2017). Timing effect on training-session rating of perceived exertion in top-class soccer referees. *International Journal of Sports Physiology and Performance, 12*(9), 1157-1162.
8. Castellano, J., Casamichana, D., Calleja-González, J., San Román, J., & Ostojic, S. M. (2011). Reliability and accuracy of 10 Hz GPS devices for short-distance exercise. *Journal of Sports Science & Medicine, 10*(1), 233.
9. Chen, M. J., Fan, X., & Moe, S. T. (2002). Criterion-related validity of the Borg ratings of perceived exertion scale in healthy individuals: a meta-analysis. *Journal of Sports Sciences, 20*(11), 873-899.
10. Chen, M., Morcillo, J. A., Rodríguez-Hernández, M. L., Zapardiel, J. C., Owen, A., & Lozano, D. (2021). The effect of weekly training load across a competitive microcycle on contextual variables in professional soccer. *International Journal of Environmental Research and Public Health, 18*(10), 5091.
11. Curtis, R. M., Huggins, R. A., Looney, D. P., West, C. A., Fortunati, A., Fontaine, G. J., & Casa, D. J. (2018). Match demands of National Collegiate Athletic Association division I men's soccer. *The Journal of Strength & Conditioning Research, 32*(10), 2907-2917.
12. Deprez D, Fransen J, Boone J, Lenoir M, Philippaerts R, Vaeyens R. Characteristics of high-level youth soccer players: variation by playing position. *J Sports Sci, 33*(3), 243–54.
13. Di Salvo V, Benito PJ, Calderon FJ, Di Salvo M, Pigozzi F. Activity profile of elite goalkeepers during football match-play. *J Sports Med Phys Fitness, 48*(4), 443–6.
14. Djaoui, L., Haddad, M., Chamari, K., & Dellal, A. (2017). Monitoring training load and fatigue in soccer players with physiological markers. *Physiology & behavior, 181*, 86-94.
15. Djaoui, L., Chamari, K., Owen, A. L., & Dellal, A. (2017). Maximal sprinting speed of elite soccer players during training and matches. *The Journal of Strength & Conditioning Research, 31*(6), 1509-1517.

16. Ehrmann, F. E., Duncan, C. S., Sindhusake, D., Franzsen, W. N., & Greene, D. A. (2016). GPS and injury prevention in professional soccer. *The Journal of Strength & Conditioning Research, 30*(2), 360-367.
17. Ekkekakis, P., Hall, E. E., & Petruzzello, S. J. (2005). Variation and homogeneity in affective responses to physical activity of varying intensities: an alternative perspective on dose–response based on evolutionary considerations. *Journal of Sports Sciences, 23*(5), 477-500.
18. Finnoff, J. T., Newcomer, K., & Laskowski, E. R. (2002). A valid and reliable method for measuring the kicking accuracy of soccer players. *Journal of Science and Medicine in Sport, 5*(4), 348-353.
19. Foster, C., Flouhaug, J. A., Franklin, J., Gottschall, L., Hrovatin, L., & Parker, S. (2001). A new approach to monitoring exercise training. *Journal of Strength and Conditioning Research, 15*, 109–115.
20. Gheller, R. G., Dal Pupo, J., Ache-Dias, J., Detanico, D., Padulo, J., & dos Santos, S. G. (2015). Effect of different knee starting angles on intersegmental coordination and performance in vertical jumps. *Human movement science, 42*, 71-80.
21. Gil, S. M., Gil, J., Ruiz, F., Irazusta, A., & Irazusta, J. (2007). Physiological and anthropometric characteristics of young soccer players according to their playing position: relevance for the selection process. *The Journal of Strength & Conditioning Research, 21*(2), 438-445.
22. Gioftsidou, A., Ispirlidis, I., Pafis, G., Malliou, P., Bikos, C., & Godolias, G. (2008). Isokinetic strength training program for muscular imbalances in professional soccer players. *Sport Sciences for Health, 2*(3), 101-105.
23. Gréhaigne, J. F., Mahut, B., & Fernandez, A. (2001). Qualitative observation tools to analyse soccer. *International Journal of Performance Analysis in Sport, 1*(1), 52-61.
24. Hader, K., Rumpf, M. C., Hertzog, M., Kilduff, L. P., Girard, O., & Silva, J. R. (2019). Monitoring the athlete match response: Can external load variables predict post-match acute and residual fatigue in soccer? A systematic review with meta-analysis. *Sports medicine-open, 5*(1), 1-19.
25. Haneishi, K., Fry, A. C., Moore, C. A., Schilling, B. K., Li, Y., & Fry, M. D. (2007). Cortisol and stress responses during a game and practice in female collegiate soccer players. *Journal of strength and conditioning research, 21*(2), 583.
26. Hongyou L, Gómez MA, Lago-Peñas C, Arias-Estero J, Stefani R. Match performance profiles of goalkeepers of elite football teams. *Int J Sports Sci Coach, 10*(4), 669–82.
27. Hughes, M., Bartlett, R., Bendiksen, M., & Carling, C. (2012). Moneyball and Soccer - an Analysis of the Key Performance Indicators of Elite Male Soccer Players by Position. *Journal of Human Sport and Exercise, 7*(2), 402–412. doi:10.4100/jhse.2012.72.06
28. Impellizzeri, F. M., Rampinini, E., Coutts, A. J., Sassi, A. L. D. O., & Marcora, S. M. (2004). Use of RPE-based training load in soccer. *Medicine & Science in sports & exercise, 36*(6), 1042-1047.
29. Izzo, R., & Giovannelli, M. (2017). Edwards TL Method and D_SHI (m): Intensity descriptors. *Sport Science, 10*(Supplement 1).
30. Jara, D., García-Hermoso, A., & Cofre-Bolados, C. (2019). Physical and Tactical Demands of the Goalkeeper in Football in Different Small-Sided Games. *Sensors, 19*(16), 3605. doi:10.3390/s19163605

31. Johnston R, Watsford M, Kelly S, Pine M, Spurris R. Validity and reliability of 10Hz and 15Hz GPS units for assessing athlete movement demands. *J Strength Cond Res, 28*(6), 1649-1655.
32. Kubayi, A., & LARKIN, P. (2020). Match Performance Profile of Goalkeepers during the 2019 COPA America Soccer Championship. *Medicina Dello Sport, 73*(3). doi:10.23736/s0025-7826.20.03724-2
33. Lagally, K. M., & Robertson, R. J. (2006). Construct validity of the OMNI resistance exercise scale. *The Journal of Strength & Conditioning Research, 20*(2), 252-256.
34. Leite, N. M. C., Gonçalves, B. S. V., Sampaio, A. J. D. E., & Saiz, S. J. (2013). Effects of the playing formation and game format on heart rate, rate of perceived exertion, vertical jump, individual and collective performance indicators in youth basketball training. *International SportMed Journal, 14*(3), 127-138.
35. Lovell, T. W., Sirotic, A. C., Impellizzeri, F. M., & Coutts, A. J. (2013). Factors affecting perception of effort (session rating of perceived exertion) during rugby league training. *International journal of sports physiology and performance, 8*(1), 62-69.
36. Malone, J. J., Di Michele, R., Morgans, R., Burgess, D., Morton, J. P., & Drust, B. (2015). Seasonal training-load quantification in elite English premier league soccer players. *International journal of sports physiology and performance, 10*(4), 489-497.
37. Malone, J. J., Lovell, R., Varley, M. C., & Coutts, A. J. (2017). Unpacking the black box: applications and considerations for using GPS devices in sport. *International journal of sports physiology and performance, 12*(Suppl 2), S2-18.
38. Malone, J. J., Jaspers, A., Helsen, W., Merks, B., Frencken, W. G., & Brink, M. S. (2018). Seasonal training load and wellness monitoring in a professional soccer goalkeeper. *International journal of sports physiology and performance, 13*(5), 672-675.
39. Matsukura, K., Asai, T., & Sakamoto, K. (2014). Characteristics of movement and force exerted by soccer goalkeepers during diving motion. *Procedia Engineering, 72*, 44-49.
40. McGahan, J. H., Mangan, S., Collins, K., Burns, C., & Dixon, J. (2018). Goalkeepers – the man in the middle. *International journal of sports science & coaching, 13*(3), 347-356.
41. McLean, B. D., Strack, D., Scott, D. A., Haigh, J. D., Sampson, J. A., & Smith, S. L. (2018). Impact of maximal-intensity repeated sprinting on indirect markers of skeletal muscle damage and remodelling in healthy, recreationally trained men. *European journal of sport science, 18*(1), 89-97.
42. Montes-Molina, R., & Rubio, V. J. (2019). Identifying key performance indicators in elite professional soccer: a network analysis approach. *International Journal of Performance Analysis in Sport, 19*(4), 626-638.
43. Nikolaidis, P. T., Clemente, F. M., van der Linden, C. M., Rosemann, T., & Knechtle, B. (2020). Validity and reliability of 10-Hz global positioning system to assess in-line movement and change of direction. *Frontiers in Physiology, 11*, 647.
44. Nolte, V. W., Noakes, T. D., & Van Vuuren, B. (2011). Physiological responses to a 4-day rugby union tournament. *Journal of sports sciences, 29*(6), 635-644.
45. Osgnach, C., Poser, S., Bernardini, R., Rinaldo, R., & di Prampero, P. E. (2010). Energy cost and metabolic power in elite soccer: A new match analysis approach. *Medicine & Science in Sports & Exercise, 42*(1), 170-178.
46. Owen, A. L., Forsyth, J. J., Wong, D. P., Dellal, A., & Connelly, S. P. (2014). Heart rate–based training intensity and its impact on injury incidence among elite-level professional soccer players. *Journal of strength and conditioning research, 28*(6), 1479-1485.

47. Randers, M. B., Mujika, I., Hewitt, A., Santisteban, J., Bischoff, R., Solano, R., ... & Mohr, M. (2010). Application of four different football match analysis systems: A comparative study. *Journal of Sports Sciences, 28*(2), 171-182.
48. Rampinini, E., Impellizzeri, F. M., Castagna, C., Coutts, A. J., & Wisløff, U. (2009). Technical performance during soccer matches of the Italian Serie A league: effect of fatigue and competitive level. *Journal of Science and Medicine in Sport, 12*(1), 227-233.
49. Raya-González, J., Castillo, D., García-Rubio, J., & Pino-Ortega, J. (2021). Impact of contextual factors on the activity profiles of soccer goalkeepers. *Research in Sports Medicine, 29*(1), 19-31.
50. Russell, M., Northeast, J., Atkinson, G., Shearer, D. A., & Sparkes, W. (2015). Repeated high-intensity effort activity in relation to changes in young soccer players' aerobic fitness. *Journal of sports sciences, 33*(19), 2033-2041.
51. Russell, M., Sparkes, W., Northeast, J., Cook, C. J., Love, T. D., & Bracken, R. M. (2016). Changes in acceleration and deceleration capacity throughout professional soccer match-play. *Journal of strength and conditioning research, 30*(10), 2839-2844.
52. Russell, M., Sparkes, W., Northeast, J., Kilduff, L. P., & Cook, C. J. (2015). Relationships between match activities and peak power output and Creatine Kinase responses to professional reserve team soccer match-play. *Human Movement Science, 45*, 96-101.
53. Sadowski, J., Mastalerz, A., Gromisz, W., & Mackala, K. (2018). Tactical and physiological demands of 4v4 and 6v6 small-sided games in elite youth soccer players. *Biology of Sport, 35*(4), 397-403.
54. Sarmento, H., Anguera, M. T., Pereira, A., Araújo, D., & Campaniço, J. (2018). Talent identification and development in male football: a systematic review. *Sports Medicine, 48*(4), 907-931.
55. Sassi, R. H., Dardouri, W., Yahmed, M. H., Gmada, N., & Mahfoudhi, M. E. (2009). Relative and absolute reliability of a modified agility T-test and its relationship with vertical jump and straight sprint. *Journal of strength and conditioning research, 23*(6), 1644.
56. Schelling, X., Torres, L., & Riera, J. (2011). Suitability of 10Hz GPS for assessing external load in collision sports: a comparison of GPS-devices. *Journal of Sports Science & Medicine, 10*(1), 635-646.
57. Souhail, H., Castagna, C., & Dellal, A. (2012). Heart rate responses and technical activity during small-sided games in elite soccer players: effects of pitch size and number of players. *Journal of Strength and Conditioning Research, 26*(1), 25-33.
58. Varley, M. C., Fairweather, I. H., & Aughey, R. J. (2012). Validity and reliability of GPS for measuring instantaneous velocity during acceleration, deceleration, and constant motion. *Journal of Sports Sciences, 30*(2), 121-127.
59. Vaeyens, R., Malina, R. M., Janssens, M., Van Renterghem, B., Bourgois, J., Vrijens, J., & Philippaerts, R. M. (2006). A multidisciplinary selection model for youth soccer: the Ghent Youth Soccer Project. *British journal of sports medicine, 40*(11), 928-934.
60. Vázquez-Guerrero, J., Rodríguez-Rosell, D., Sánchez-Medina, L., Rodríguez-Matoso, D., González-Badillo, J. J., & Pareja-Blanco, F. (2020). Differences in the load-velocity profile between 4 competition levels and genders in elite volleyball players. *International Journal of Sports Physiology and Performance, 15*(10), 1503-1511.

61. Waldron, M., Highton, J., & Twist, C. (2011). A comparison of the Match Demands of Elite and Semi-Elite Rugby League Competition. *Journal of Strength and Conditioning Research, 25*(11), 3156-3161. doi:10.1519/jsc.0b013e318212dea7
62. Wehbe, G. M., Hartwig, T. B., Duncan, C. S., & Mckenna, M. J. (2015). GPS Analysis of an International Field Hockey Team: A Case Study. *Journal of Strength and Conditioning Research, 29*(5), 1371-1376. doi:10.1519/jsc.0000000000000773
63. Weston, M., Siegler, J., Bahnert, A., McBrien, J., & Lovell, R. (2015). The application of differential ratings of perceived exertion to Australian Football League matches. *Journal of Science and Medicine in Sport, 18*(6), 704-708.
64. White, A., Hills, S. P., Cooke, C. B., & Batten, T. (2019). The Efficacy of the New Acceleration Points System (Aps) for the Technical and Tactical Analysis of Field Hockey. *Journal of Strength and Conditioning Research, 33*(5), 1373-1381. doi:10.1519/jsc.0000000000002519
65. Wiig, H., & Wassermann, J. D. (2021). Performance profiles in soccer: A systematic review. *International Journal of Sports Science & Coaching, 16*(1), 3-23.
66. Williams, J. D., Abt, G., & Kilding, A. E. (2010). Ball-sport endurance and sprint test (BEAST90): validity and reliability of a 90-minute soccer performance test. *Journal of Strength and Conditioning Research, 24*(12), 3209-3218.
67. Windt, J., & Gabbett, T. J. (2017). Is it all for naught? What does mathematical coupling mean for acute:chronic workload ratios? *British Journal of Sports Medicine, 51*(5), 597-598.
68. Wong, P. L., Chamari, K., Dellal, A., & Wisløff, U. (2009). Relationship between anthropometric and physiological characteristics in youth soccer players. *Journal of strength and conditioning research, 23*(4), 1204-1210.
69. Yu, J., Schmitz, R. J., & Zhang, S. (2018). Effects of soccer shoes with different properties on jumping ability and kicking ball velocity in untrained women. *Sports biomechanics, 17*(4), 467-476.
70. Zapartidis, I., Skoufas, D., & Vareltzis, I. (2007). Angles and velocities of trunk segment movements in soccer kicking. *Journal of Sports Science & Medicine, 6*(1), 19-23.
71. “White Paper: An Introduction to Inertial Movement Analysis (IMA).” Catapult,. (2018). www.catapultsports.com/blog/white-paper-introduction-ima.