ASSESSING THE ROLE OF THE FUNCTIONAL MOVEMENT SCREEN IN INJURY PREVENTION AND HOW TECHNOLOGY CAN ENHANCE ITS USE IN APPLIED SETTINGS

PAUL D SMITH

SUPERVISOR – DR MICHAEL HANLON

Department of Sport and Exercise Science, Waterford Institute of Technology, Ireland

For the award of Doctorate of Philosophy

Submitted to Waterford Institute of Technology, September 2018

Abstract

ASSESSING THE ROLE OF THE FUNCTIONAL MOVEMENT SCREEN IN INJURY PREVENTION AND HOW TECHNOLOGY CAN ENHANCE ITS USE IN APPLIED SETTINGS

P Smith, Department of Sport and Exercise Science, Waterford Institute of Technology, Ireland

The Functional Movement Screen (FMS) assesses an individual's mobility, stability and flexibility to identify any weaknesses or asymmetries that may exist. Research has considered if it can predict injury and performance, with inconclusive results. Reliability is high for total FMS scores, but decreases significantly for subtest scores. Study one assessed whether the FMS could accurately predict injury rates and performance in 116 soccer players (mean age=23.2±4.4 years). Players undertook the FMS and performance tests, and injury occurrence was tracked throughout one season. Results showed a cut-off score of 14 or less could not predict a player's risk of all injuries (odds ratio=1.01), or non-contact injuries (odds ratio=0.63), but was significantly linked to vertical jump height (p=0.006). Additional studies discussed the development and validation of novel software to automatically assess the FMS. Study two assessed the software's validity and test-retest reliability when scoring the deep squat (DS). Two validation sessions were completed. Initially, 141 participants (mean age=9.7±3.7 years) performed the DS three times, before 16 participants (mean age=29.8±8.1 years) performed the DS on two occasions, 72 hours apart. Results showed the software had good agreement (87.5%) and moderate correlation (r_s=0.44) with manual scoring and had moderate test-retest reliability (ICC=0.74). Study three developed the software to assess three FMS subtests. Validation involved 27 participants (mean age=19.8± 3.8 years) performing each subtest three times, with the software showing good mean agreement (87.3%), and inter-rater reliability (K_w=0.76), and strong correlation (r_s=0.76), with live manual scoring. The final study assessed the software's validity and test-retest reliability compared to manual scoring for all seven subtests. Twenty-three participants (mean age=22.8±5.3 years) completed all seven FMS subtests. The software's mean agreement (87.5%) with live and video manual scoring was good, and a strong mean correlation was also reported (r_s=0.71), for the seven subtests. Mean test-retest reliability was good across all seven FMS subtests (ICC=0.96). Overall, this body of research suggests the FMS cannot accurately be used to predict injury or performance in soccer players, but the development of novel software has provided a valid alternative to manual assessment for certain FMS subtests, which could provide benefits for its use in applied settings.

Author's Declaration

I hereby declare that this thesis does not contain any material that has been previously published or been submitted to any other third level institution or for any higher degree in this institution. The source of the information contained within this thesis is solely the work of the author except where specified.

Name:

Signature:

Date:

Acknowledgements

Thank you to the various clubs that gave me valuable access to their players, training facilities and medical staff and to all the participants involved across this programme of research. Thank you also to my supervisor, Dr Michael Hanlon, whose continued support, advice and guidance over the last four years has helped me take my initial ideas and thoughts forward into a meaningful, interesting and enjoyable research project. Lastly, and most importantly, a big thank you to my amazing wife, Natasha, and my two wonderful children, Holly Mae and James, for their patience and love throughout the project. I hope they will see the rewards of that support in years to come.

Table of Contents

CHAPTER ONE

CHAPTER TWO

CHAPTER THREE

THE EFFECTIVENESS OF THE FUNCTIONAL MOVEMENT SCREEN IN PREDICTING INJURY RATES IN SOCCER PLAYERS 73

CHAPTER SIX

CHAPTER SEVEN

APPENDIX A

APPENDIX B

APPENDIX C

APPENDIX D

APPENDIX E

APPENDIX F

APPENDIX G

List of Figures

List of Tables

CHAPTER ONE

THESIS INTRODUCTION

1.0. INTRODUCTION

Injury rates in soccer have been an issue for several years as evidenced by the number of training and match days players miss each season (Ekstrand, Hägglund, & Waldén, 2009). Hawkins & Fuller (1999) reported that the overall risk of injury to professional soccer players is approximately 1000 times higher per 1000 hours, than for industrial occupations generally regarded as high risk. In a study of 23 of the top 50 teams in Europe as selected by the Union of European Football Associations (UEFA) from 2001 to 2008, 4,483 injuries occurred during 550,000 hours of exposure (Ekstrand, Hägglund, & Waldén, 2009). Results showed a squad of 25 players can expect to receive 50 injuries per season. Injuries during matches were more common than in training situations and injury incidence increased with time in each half. During an audit of injuries in professional soccer over two seasons, Woods et al. (2004) reported hamstring strains to be the most common injury amongst 91 of the 92 soccer clubs from the English football leagues.

Ekstrand et al. (2009) found that the injury incidence from 2001 to 2008 remained stable with no significant differences reported between seasons. When considering additional research undertaken in this area looking at all types of injuries, there appears to be no significant decrease in injury incidence over the past ten years. In fact, Junge & Dvorak (2013) reported a slight increase in the number of injuries per match over time for females playing in the Federation International de Football Association (FIFA) World Cups and Olympic Games soccer tournaments from 1999 to 2011. Given the advancements in education and treatment relating to injuries, a reduction in all injury incidence for both male and female players could be expected, yet has not been seen across all injury types and ability levels.

With injuries in soccer still occurring frequently, there is on-going research in injury prevention techniques that could help reduce the number of injuries recorded. Consideration has been given to warm ups (Marshall, Lopatina, Lacny, & Emery, 2016; Thompson et al., 2016; Towlson, Midgley, & Lovell, 2013), recovery techniques (Higgins, Greene, & Baker, 2017), training load monitoring (Bacon & Mauger, 2017; Halson, 2014), flexibility (Rey, Padron-Cabo, Barcala-Furelos, & Mecias-Calvo, 2016; Witvrouw, Danneels, Asselman, Have, & Cambier, 2003), strength and conditioning programmes (Heiderscheit, Sherry, Silder, Chumanov & Thelen, 2010; Van Beijsterveldt, Van Der

Horst, Van De Port, & Backx, 2013), nutrition (Beck, Thomson, Swift, & von Hurst, 2015) and other areas (Munro, Herrington, & Comfort, 2012), in an attempt to identify techniques that could positively impact injury rates. To date there has been limited research that has considered whether movement screening can effectively be used to assess injury risk within a population of soccer players, whilst taking into account match and training exposure levels, individual subtest scores and number of asymmetries. Whilst injuries are multi-factorial in their nature (Hägglund, Waldén, & Ekstrand, 2013), making them very difficult to predict, movement screening could be a useful tool to indicate potential injury risk within this population type.

Cook, Burton, & Hoogenboom (2006) suggest the main goal of movement screening is to decrease injuries, enhance performance and improve overall quality of life. Movement screens are conducted to assess an individual's fundamental movement patterns, which can include mobility, stability, balance and flexibility. Screens help a health professional to identify weaknesses or limitations that exist which can be linked to a reduction in that individual's performance as a direct result of their compensatory movement patterns. Asymmetries, a difference between the right and left-hand side of the body in regard to fundamental movement, can also be established through this process. The results of such screens are typically used by the health professional to develop a corrective exercise programme for the individual, with the goal of improving their fundamental movement patterns leading to reduced risk of injury (Cook et al., 2006).

One specific seven subtest Functional Movement Screen (trademarked as the FMS) has gained particular popularity. These seven subtests were developed by Cook et al. (2006) in an attempt to fill the void between pre-participation screens and performance tests by evaluating individuals in a dynamic and functional setting. The screen involves seven fundamental movement patterns that require a combination of stability and mobility. They are designed to provide observable and measurable performance of basic stabilizing movements and place individuals in extreme positions where weaknesses and imbalances can be identified if stability and mobility is not utilized (Cook et al., 2006).

The FMS has been the subject of increased levels of research in the past ten years, which has discussed its reliability and its ability to predict injury rates and performance amongst certain populations. It has been suggested that an athlete's performance on the FMS can help to predict their risk of injury. Kiesel, Plisky, & Voight (2007) suggested a score of 14 out of a possible 21 as being the cut-off point for possible injury, in their study of 46 NFL American footballers over one season. Further research on 38 Division II Collegiate athletes stated that 11 out of 16 participants within their sample who scored 14 or less on the FMS sustained an injury during one season of injury rate tracking, compared to 8 out of 22 from the group scoring ≥15 (Chorba, Chorba, Bouillon, Overmyer, & Landis, 2010). This suggests the FMS can help to identify those at risk of injury within various sporting populations. However, studies that have reported a link between FMS score and injury have had several limitations including; the injury definition used, small sample sizes, undefined exclusion criteria and a lack of reference to exposure levels. This makes it difficult to assess how effective the FMS is in predicting injury rates amongst different populations.

It has also been found that the FMS has moderate to excellent inter-rater reliability, when considering only the total FMS composite score, among novice raters (Teyhen et al., 2012). Minick, Kiesel, Burton, Taylor, Plisky, & Butler (2010) reported that the total FMS composite score could confidently be used to assess the movement patterns of athletes and to make decisions when prescribing an intervention for performance. These studies suggest that the total FMS composite score can be used reliably in an applied setting, even by testers with minimal experience of movement screening.

As the FMS uses an ordinal scale between zero and three to score each subtest, resulting in a maximum total composite score of only 21, it is perhaps not that surprising that minimal differences have been found in total FMS composite scores between raters of mixed experience when scoring FMS performance of the same individual. Therefore, the test-retest reliability of the FMS and the inter-rater reliability of each FMS subtest is potentially of more interest.

Test-retest reliability of the total FMS composite score has been shown to be weaker compared to inter-rater reliability for live scoring (Teyhen et al., 2012), although agreement remains excellent for video recordings (Parenteau et al., 2014; Teyhen et al., 2012). It therefore appears that test-retest reliability is high when using video recording but reduces for live scoring when considering total FMS composite scores. When assessing inter-rater reliability of subtest scores, the agreement levels decrease significantly compared to total FMS composite scores (Gulgin & Hoogenboom, 2014; Minick et al., 2010; Shultz, Anderson, Matheson, Marcello, & Besier, 2013). These studies indicate that certain subtests cannot be scored as reliably as others by different raters, and that the relatively simplistic scoring system used in the FMS can in fact contribute to this reduced reliability, particularly when considering mid-range performance.

This thesis initially explored the use of the FMS and its links with performance and injury. The first study examined its effectiveness at predicting injury rates amongst a sample population of League of Ireland soccer players, through undertaking a battery of screening tests and tracking injury rates over one season. To build upon existing research in this area, consideration was given to match exposure levels, subtest scores and number of asymmetries, and how these impact injury risk, in addition to the total FMS composite score. A comparison of FMS score and performance test results was also conducted to assess the most effective method by which to assess injury risk. Following completion of study one, studies two, three and four focussed on the development and validation of new motion tracking software designed to reduce the subjectivity of FMS scoring by providing a reliable and valid alternative to manual scoring. Study two discusses the development of prototype software designed to automatically assess the FMS deep squat subtest and the process of validating same to compare its validity against manual scoring and the test-retest reliability of the software. Study three built upon the work completed in study two by further developing the software to allow it to assess performance on three FMS subtests. The software’s automated scoring system was validated to understand whether it offered a more reliable and valid method by which to assess the FMS when compared to manual live scoring. Finally, study four discussed the development and validation of motion tracking technology to assess all seven FMS subtests and its ability to prescribe corrective exercise programmes compared to two certified FMS testers. The development and validation of such software could reduce the subjectivity of FMS scoring in an applied setting, which could lead to more suitable corrective exercise programmes being prescribed to individuals following completion of the FMS (Figure 1.0).

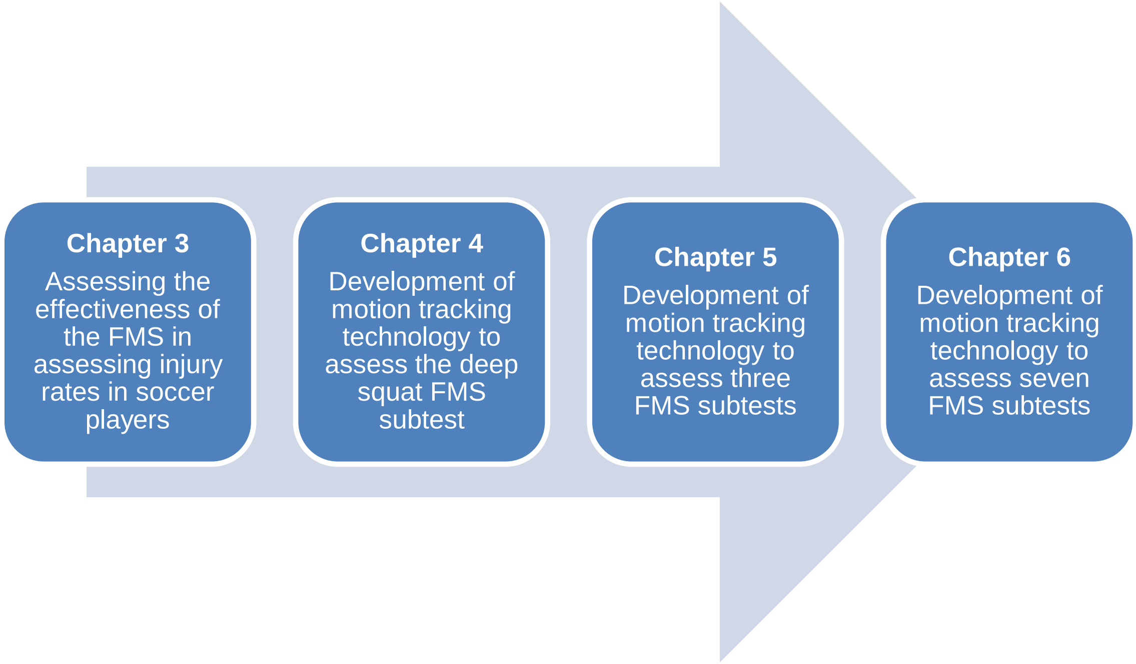


Figure 1.0. Overview of programme of primary research

CHAPTER TWO
LITERATURE REVIEW

2.0 INTRODUCTION

This literature review covered the most common type and location of injuries in soccer, the time loss from play resulting from these injuries and the main risk factors attributed to causing such injuries. This information provided a clear picture of the scale of the issue of injuries in soccer across various levels of the game. Tracking the level of injuries is an important area of study for researchers and practitioners, as they attempt to increase player availability, and thus there has been several injury surveillance audits conducted within a soccer environment in the last 20 years. In this literature review, consideration was given to the findings of these studies as well as the procedures and criteria that were used to measure and monitor injury occurrence, which included aspects such as how injuries were reported, the length of the research and the injury definitions used within each study.

This review also considered the use of movement screening and research studies that have been conducted analysing links with injury and performance. Movement screening is an activity employed by health professionals to measure and evaluate a person's fundamental movement patterns to diagnose any weaknesses or compensations that may exist. A body of research has been undertaken in this area that discusses the most reliable protocols to follow when assessing movement, the rising popularity of the Functional Movement Screen (FMS), the reliability of the FMS and its ability to predict injury and performance in certain populations. This review considered these research papers and provided critical insight into the strengths and weaknesses of same, to provide the reader with a clearer understanding of this field of work.

Lastly this review discussed the use of technology to evaluate movement characteristics and biomechanical variables, and how these are used in an applied setting to enhance sports performance and reduce injury risk. Consideration was given to marker based motion capture systems and inertial measurement units and the validity and reliability of such methods to inform practice. Finally, the review considered how the emergence of new markerless, depth camera technology, could provide a more accessible and cost-effective solution compared to current gold standards within the sports science industry.

2.1 INJURIES IN SOCCER

2.1.1 INJURY TYPE

One of the most comprehensive injury surveillance studies conducted to date, within soccer, is that by Ekstrand, Hägglund, & Waldén (2009) which tracked injury rates amongst the top 50 professional teams across Europe, as defined by UEFA, for seven years. This study reported that 87% of injuries recorded affected the lower extremities in their study of 23 professional clubs. When considering the type of injury, they found muscle strains (35% of all injuries recorded), ligament sprains (18%) and contusions (17%) to be the most common. Earlier research reported similar findings with the main injury types being muscle strains (37%), ligament sprains (19%) and contusions (7%) in an extensive study involving 2376 professional players from the English football leagues over two seasons (Hawkins, Hulse, Wilkinson, Hodson, & Gibson, 2001). Ekstrand (2008), in a smaller study of 17 UEFA Champions League teams conducted over a five-season period, and 14 Swedish Super league teams over two seasons, reported over 6300 injuries and found overuse injuries to be the most commonly reported, comprising almost one third of injuries, followed by muscle strains and ligament sprains. Research has shown that non-contact injuries appear to be more common than contact injuries, with Hawkins et al. (2001) finding that 38% of all injuries reported (n=6030) in their study were caused due to contact between players. In a study of 266 professional players from eleven European clubs, it was reported that 23% of all injuries recorded (n=360) were the result of foul play, and all of them due to an opponent foul (Waldén, Hägglund, & Ekstrand, 2005).

The most common injury types remain constant when considering specific periods of a season. Woods, Hawkins, Hulse, & Hodson (2002) reported muscle strains (37%), ligament sprains (19%) and contusions (7%) to be the most common injury type from 1025 injuries recorded during pre-season periods over two seasons from a study involving a large cohort of 2376 players. This study also showed contact injuries to be less common than non-contact, with 29% of all injuries recorded being caused by contact. Studies of tournament soccer over a period of four to six weeks have shown that the three main injury types remain constant with that previously reported, but the most common injury was no longer the muscle strain. In their comprehensive study, Junge & Dvorak (2013) monitored injury rates in FIFA tournaments and Olympic Games from 1998 to 2012. They reported a total of 3944 injuries from 1546 matches, which included male, female, u20, u19 and u17 tournaments. Contusions (55%), sprains (15%), and muscle strains (10%) were reported as the three most commonly occurring injuries across all tournaments. It was also reported that contact injuries accounted for 80% of all injuries recorded, which is significantly higher than in any other research, and that 47% of these contact injuries were the result of foul play.

Although the main injury types reported are the same as in other studies, there is a significant difference between the number of contusions reported by Junge & Dovark (2013) compared to others (55% to just 7% reported by Woods et al. 2004). There are also a much higher percentage of contact injuries (80%) than in any other study. One reason for this could be the greater mix of age groups in the study and that injuries were also recorded for female tournaments as well as male. Woods et al. (2002) only included male professional soccer players in their study. The demands of the game for female and youth soccer does vary considerably from that of the male elite game which in turn could lead to the differences seen in types of injury reported. Additionally, the definitions of injury between these two studies differed, with Junge & Dvorak (2013) defining an injury as any physical complaint incurred during a match that received medical attention. Woods et al. (2002) used a time loss definition and only recorded injuries that resulted in absence from participation in training and match play for two days or more. Contusions would often lead to a physical complaint during a match, which although painful, may not necessarily result in any actual time loss as the player continues to play on in the match. As contusions and contact injuries are more likely to occur during matches this could explain the increased level of this type of injury reported by Junge & Dvorak (2013). The differences reported between these two studies regarding injury types, highlight the importance of understanding the injury definition employed when examining and comparing results between studies, to draw accurate conclusions and use them in applied settings.

From the body of research that has been undertaken it is possible to conclude that the most common injury types for soccer players are contusions, muscles strains and ligament sprains. When using a time loss definition of injury, muscle strains and ligament sprains become even more prominent. This trend appears to remain constant no matter whether the sample population is male, female, u19 age group, u17 age group or participating in soccer at different stages of the calendar year. Variations were seen in regard to the percentage of each injury type across research papers, but these can largely be attributed to the way injuries were defined, the sample sizes used and the ability/age differences across samples.

2.1.2 INJURY LOCATION

In a study of 23 professional soccer clubs, the thigh, knee, ankle and hip/groin were the most common injury locations (Ekstrand et al., 2009). Thigh injuries (23%) were more common than knee (18%), ankle (14%) and hip/groin injuries (14%) over the seven seasons of the study. Hawkins et al. (2001) reported similar findings, with the thigh (23%), knee (17%), ankle (17%), and lower leg (12%) being the locations where most injuries occurred over two seasons. Similarly, Hägglund, Waldén, & Ekstrand (2006) reported that the thigh (22.6%), hip/groin (17.3%) and the knee (16.82%) were the most common locations for injury from the 1,189 reported over two seasons in their study of 263 male elite Swedish soccer players. These three studies show clear similarities regarding the most common injury locations, and perhaps more strikingly, similar percentages of each injury location, despite vast differences in sample sizes and total injuries recorded.

As with injury type, the three most common injury locations appear to be quite consistent no matter the time of season. Woods et al. (2002) reported that the thigh (23%), ankle (17%), and knee (16%) were the most common locations of injuries reported during a pre-season period of July to August. This differs slightly when considering female instead of male participants. In their study between March and November, Nilstad et al. (2014) found that the knee (31%) was the most commonly reported injury in their sample of female soccer players, followed by the ankle (23%) and then the thigh (21%). This highlights the higher risk of knee injuries that female player's face, which has been widely researched in recent times to reduce the incidence of such injuries in female players (Herrington, Munro, & Comfort, 2015; Jones, Herrington, & Graham-Smith, 2015).

The thigh, ankle, knee, and hip/groin are the most common location of injuries within soccer based on research conducted. There is variation in the incidence levels of each of these locations, yet these are undoubtedly the four main locations of injury when using the "time loss" method of defining injury for both male and female players. Given the nature of the sport and its high intensity, multi directional and explosive demands, the variation of occurrence amongst the four locations is not surprising, particularly when allowing for risk factors that are specific to each player every time they play. Additional contact-based injuries located around the head and neck are also common when defining injury as any incidence that requires medical treatment during a match situation.

2.1.3 INJURY SEVERITY

As the previous two sections have outlined, injury incidence is high amongst soccer players. It is therefore important to understand the severity of the injuries that are being recorded and the impact this has on the amount of time players miss from participating in soccer activities. One Premier League team alone lost a combined 2289 days to injury across their squad during the 2016/17 season compared to the average number of combined days per team being 1,410 (Physioroom.com, 2017). Considering an average season is around 300 days including pre-season, this highlights the severity of injuries to elite players that are occurring and the impact this can have on their team's chances of success each year.

The UEFA study (Hägglund, Waldén, Bahr, & Ekstrand, 2005) classified injury severity as follows:

- Slight – 1 to 3 days absence from soccer participation
- Minor – 4 to 7 days absence
- Moderate – 8 to 28 days absence
- Severe - >28 days absence

Absences were based on actual days missed rather than estimates made by the team's medical physician.

Table 2.1. Summary of the severity of injuries recorded in previous injury surveillance studies in soccer (Highest value in each row in bold)

Authors and Date	Total Number of Injuries	Slight Injuries 1 - 3 days	Minor Injuries 4 - 7 days	Moderate Injuries 8 - 28 days	Severe Injuries >28 days
Hawkins et al., 2001	6030	588 (10%)	1385 (23%)	**2698 (45%)**	1359 (23%)
Woods et al., 2002	1025	130 (13%)	274 (27%)	**420 (41%)**	201 (20%)
Hagglund et al., 2006	1189	**394 (33%)**	328 (28%)	342 (29%)	125 (10%)
Ekstrand et al., 2009	4483	971 (22%)	1164 (26%)	**1651 (37%)**	697 (15%)
Nilstad et al., 2014	171	24 (14%)	32 (19%)	**64 (37%)**	51 (30%)

Table 2.1 highlights the different severities of injuries recorded in five previous injury surveillance studies completed on soccer players. It clearly identifies some inconsistencies that can skew the injury data recorded and lead to inaccurate conclusions being drawn. For example, slight injuries, as defined by the UEFA study, reduce in percentage terms when removing single day absences from soccer activity as shown in the first two papers in Table 2.1. Both studies defined an injury as one which prevented the injured player from participating in normal training or competition for more than 48 hours. As such, an injury that resulted in a player missing a training session the day after the injury occurred would not have been recorded in these studies. Moderate injuries (8 to 28 days absence) are the most commonly reported injuries in four of the five studies highlighted in Table 2.1. Outside of this trend there is no consistency in terms of the most commonly reported injury severity period across the studies. Ekstrand et al. (2009) in their comprehensive study of 2226 professional players across Europe over seven seasons from 2001 to 2008, reported that the incidence of severe injuries did not differ significantly between seasons. This is somewhat surprising given the advances in injury prevention techniques that have become common place across all levels of the sport over the last ten years.

2.1.4 RISK FACTORS FOR INJURY

It would be useful to understand the main risk factors that leave players more predisposed to injury than others. Understanding such risk factors should assist practitioners when developing training programmes and interventions to help prevent such injuries occurring in the first place.

Age

It has been suggested that the age of a player can affect their likelihood of receiving an injury. It is assumed that the older a player gets the more at risk they become. In a study of 20 male soccer teams (306 players) from the Icelandic elite and first divisions, Arnason et al. (2004) found that the injured group of players were significantly older than the non-injured group of players and as such older players were more at risk of injury in general (an increase in odds ratio of 0.1 per year, $p=0.05$). These authors also suggested that older players were more at risk of hamstring injuries in particular. Hägglund et al. (2006) also reported that increasing age was a significant risk factor for hamstring and ankle injury. However, they did not suggest an association between increased age and injury risk in general. In a more recent yet smaller study of 36 professional players from the English Premier League over one season, it was reported that with each additional year of age, the odds of sustaining an injury increase by x1.78 (Henderson, Barnes, & Portas, 2010). However, Nilstad et al. (2014) reported no association between age and the risk of lower extremity injuries in general. As this study involved elite female players the impact of age on injury occurrence for this cohort could differ from that seen for male players and is worthy of further research. The existing research suggests that age could affect injury risk within soccer, yet further work needs to be conducted within this area before a definitive relationship between age and injury can be determined.

Fatigue

Ekstrand et al. (2009) reported that the incidence of match injuries in their study showed an increasing frequency over time in both the first and second halves. It could be suggested that fatigue played a role in these injuries occurring as studies have shown that fatigue is increased towards the end of a game (Mohr, Krustrup, & Bangsbo, 2005; Krustrup, Sebis, Jensen & Mohr, 2010). This theory is supported by Hawkins et al. (2001) who reported that a greater than average frequency of injuries in their study were observed during the final 15 minutes of the first half and the final 30 minutes of the second half, with the final 15 minutes of the match being the period when most injuries occurred (25%). They also found that more injuries were reported in the second half compared to the first half (57% to 43%). Research by Greig & Siegler (2009) on 10 male soccer professional players, analysing the effect of soccer specific fatigue on peak eccentric torque of the knee flexor muscles, identified that eccentric hamstring strength decreased as a function of time and immediately after the half-time interval. This suggests an increased risk of injury within these muscles, during the second half, especially when completing explosive movements. Additionally, these authors reported that peak hamstring torque was lower at the end of the half-time period compared to the start of the half-time period, highlighting the importance of a re-warm up strategy prior to the start of the second half to help reduce injury risk. An additional study by Small, McNaughton, Greig, & Lovell (2010) involving 16 semi-professional soccer players, also reported a significant decrease in eccentric hamstring peak torque during a 90-minute simulated match situation (16.8% decrease, $p<0.01$). They also found a significant decrease in functional eccentric hamstring to quadriceps ratio over the same 90-minute period (15% decrease, $p<0.01$). Such decreases would result in an increased risk of injury within the hamstring muscles, which could be caused by fatigue. These studies had small sample sizes, primarily due to the difficulties involved in recruiting large numbers to participate in a simulated 90-minute match and pre-test/post-test protocol due to the busy schedule faced by elite level teams. However, the results suggest that an injury prevention programme and a pre-second half re-warm up routine, would be beneficial in reducing the impact of fatigue on injury occurrence.

Understanding when a player is fatigued during training and match situations would help practitioners to assess their current injury risk in real time. One such method that has gained popularity is the use of Global Positioning System (GPS) technology. These are typically worn to assess levels of high intensity running, total distance covered, number of accelerations/changes of directions and various other movements by each player, providing an indication of the load placed upon each player in real time. However, the expensive nature of these systems means the use of GPS monitors is very much reserved for elite level teams. Question marks have also been raised in regard to the reliability of such systems when measuring high intensity movements such as sprinting (Gray et al., 2010). GPS technology has recently been approved for use during competitive soccer matches by FIFA. However, this ruling has not been adopted by all the professional leagues across the globe, and as such research in this area is still in its infancy. This technology has however, been available for use in other professional sports such as Rugby Union and League during matches for a number of years, Research in this area has highlighted the different running demands placed upon backs and forwards during an 80-minute match (Dempsey, Gibson, Sykes, Pryjmachuk, & Turner, 2017), yet these demands were eliminated when reported relative to match minutes played. When considering fatigue, research has reported that the relative distance covered at high velocities and the changes in velocities by 19 Rugby Sevens international level players were reduced between 1 and 16% from the first to the second half (Higham, Pyne, Anson, & Eddy, 2012), but there was limited evidence of these players suffering accumulated fatigue across multiple games in tournament situations. The same study also found that substitutes playing an average of 3.01 minutes per game, exhibited a significantly higher work rate compared to players who had played the entire match. From a soccer perspective, the ability to analyse GPS data collected from matches should greatly aid practitioners understanding of the load placed upon individual players and how and when fatigue starts to increase their risk of injury, so they can tailor strength and conditioning programmes accordingly.

At the elite level, such as the English Premier League, computer aided motion capture systems have been developed to track player’s movements throughout a 90-minute competitive match. Using such systems to simply compare distance covered during first and second half periods is however not an effective means by which to measure levels of fatigue (Carling, Bloomfield, Nelsen, & Reilly, 2008). This is because the tactics employed, current score line and each player’s individual effort levels all impact the distance covered at any particular point in a match, and as such distances covered in each half vary from game to game due to factors other than fatigue. Analysing players overall speed and/or the sprint distances and speeds in the second half compared to the first half is perhaps a more reliable manner by which to measure fatigue using motion capture analysis (Carling et al., 2008). However, factors such as the current score line have been shown to significantly affect the number of high intensity actions completed by players (O’Donoghue & Tenga, 2001), and as such the exact role fatigue plays in reducing sprint speed and or distance is difficult to establish. Carling et al. (2008) suggested that further studies relating to work rate over several games and injury occurrence would be a useful way to identify the effect of fatigue on injury risk in match situations. Playing position could also impact levels of fatigue as the demands placed upon players in a match differ depending on position. Bloomfield, Polman, & O’Donoghue (2007) reported that playing position significantly influences the percentage of purposeful movement time spent completing activities such as running, sprinting and shuffling during their study of 55 Premier League players.

Fatigue can be determined as a significant risk factor for injury within soccer. The second half appears to provide increased risk compared to the first, with the final fifteen minutes of the second half of particular high risk. This trend has remained constant over the past fifteen years and would suggest further work needs to be undertaken in regard to preparing players for the demands of the game of soccer if injuries are to be reduced. Further consideration should also be given to how best to interpret the data from time motion analysis systems during matches, to understand individual levels of fatigue at any given point in a game. Fully understanding this real-time data, and the role fatigue plays for each individual player, could significantly impact the role substitutions play in not only improving team performance, but also helping to reduce injuries by removing certain players from the field of play before they reach a level of fatigue that puts them most at risk of injury.

Flexibility

Consideration has been given to the role flexibility can play in injury occurrence. It has been suggested that players that are less flexible, particularly within their lower extremities, are more at risk of injury. Arnason et al. (2004) conducted four flexibility tests on 306 players to measure hamstring, hip, groin and quadriceps flexibility. They reported a trend of lower hip adductor flexibility within those players that suffered groin strains (n=13) compared to those that didn't (n=485). In a smaller study of 146 male professional players from the Belgian league over one season, it was found that a significant correlation existed between players with decreased hamstring flexibility and the occurrence of a hamstring muscle injury (Witvrouw et al., 2003). In the same study, a similar, but weaker, relationship was found for the quadriceps muscles. Both studies used the same flexibility test for the quadriceps muscles yet only Witvrouw et al. (2003) reported any relationship between injury and quadriceps flexibility. A slightly different hamstring flexibility test was administered by Witvrouw et al. (2003), to that used by Arnason et al. (2004). Both studies reported a link between hamstring flexibility and hamstring injury. Using the same passive straight leg raise test as Witvrouw et al. (2003) more recent research with a small cohort of only 36 Premier League players identified a similar trend between hamstring flexibility and an increased risk of injury (Henderson, Barnes, & Portas, 2010). These similar results occurred despite contrasting definitions of injury being used.

Further research relating to soccer injuries has found no correlation between lower extremity muscle flexibility and increased risk of injury. Söderman, Alfredson, Pietilä, & Werner (2001) reported that hamstring flexibility was not associated with traumatic leg injuries in their study of 146 female players. This was further emphasized by Rolls & George (2004) in their study of 111 players aged 9 to 19 attached to a professional football club. They found no significant correlation between hamstring length, as measured by flexibility tests, and injury occurrence when analysing the 16 hamstring injuries reported during the study. Although these two studies included cohorts of differing age and gender, and implemented different flexibility tests, neither reported a significant correlation between flexibility and hamstring injury. When considering other sports, a small study involving 126 community level Australian Rules Footballers over one season reported that hamstring flexibility was not significantly related to hamstring injury when tested using the active knee extension and passive straight leg raise tests (Gabbe, Finch, Bennell, & Wajswelner, 2005). Interestingly, this study did report a link between the group of players with the lowest quadriceps flexibility and a decreased probability for completing a season without sustaining a hamstring injury, when using the Modified Thomas test to measure quadriceps flexibility. Yeung, Suen, & Yeung (2009) similarly found that hamstring flexibility was not related to hamstring injury when using a straight leg raising test to measure flexibility in their study of 44 sprinters from the Hong Kong Sports institute over a 12-month period in their prospective study.

As outlined in the previous two paragraphs, one of the difficulties regarding analysing the role flexibility can play in causing injury, are the tests used to measure flexibility, their reliability and how they are carried out by the researchers. The use of different tests across the various studies makes it difficult to draw any significant conclusions relating to flexibility and injury occurrence or risk. The occurrence of hamstring injuries during high speed running occurs most frequently during the terminal swing phase of the gait cycle, whereby the hamstring lengthens and absorbs energy prior to foot contact. During kicking motions, hamstring injuries occur during slow or fast movements that involve simultaneous hip flexion and knee extension, which place the muscle in a position of extreme stretch (Heiderscheit et al., 2010). In both instances, increased flexibility within the hamstring muscles, would reduce the strain placed on the muscles when in a stretched position, which in theory should reduce the risk of injury. However, to date the research outlined in previous paragraphs does not support the theory that increased flexibility reduces injury risk. Further research that standardises testing protocols and considers aspects such as hamstring strength alongside flexibility, should be considered to improve knowledge and understanding in this area.

Previous injury

Previous injury has been clearly identified as an important risk factor for injury in soccer players. Hagglund et al. (2006) found that 87% of players injured in one season received an injury the following season compared to only 48% of players that were not injured in the first season. They reported a three-fold increase in risk of suffering an injury the season after first suffering an injury. This high level of increased risk could have been influenced by the length of the study and the researcher's ability to track injury occurrence compared to other studies. Often an injury is only considered a re-injury if it occurs within three months of that players return to training and is in the same site as the previous injury. In the study by Hagglund et al. (2006) any injury that occurred in the second season of the study was considered. This highlights the importance of study length in accurately establishing risk factors for injury. In their study of 508 Norwegian Division One, Two and Three players, Engebretsen, Myklebust, Holme, Engebretsen & Bahr (2010) reported that previous acute hamstring injury was a significant risk factor for new hamstring injuries and that previously injured players have more than twice as high a risk of sustaining a new hamstring injury. This study was conducted over one pre-season period and with lower level teams than other studies. The medical support available, and rehabilitation programmes implemented, to amateur teams could be a factor in the re-injury rates seen within this study, as they are unlikely to have been to the standard seen within a professional team set up. Additionally, the re-injury rates could be higher in pre-season due to the players showing lower levels of fitness and hence fatigue playing a larger role. It is widely acknowledged that amateur teams are not as match fit as professional teams, which could also impact re-injury rates.

Other studies have also concluded that re-injury is a major risk factor for injury occurrence. Ekstrand et al. (2009) found that re-injuries constituted 12% of all 4483 injuries recorded and caused a significantly longer absence than non-re-injuries (24 v 18 days). Woods et al. (2004) found that the re-injury rate for the hamstring was 12% over two seasons, yet certain clubs within their study reported no recurrence of hamstring injuries over that period. This suggests that certain variables play a major part in preventing re-injury such as medical diagnosis, training techniques and medical management. This highlights the important role support staff play in injury prevention, and particularly re-injury occurrence. Ekstrand (2008) suggests that controlled rehabilitation including rules for return to team training and matches help reduce the risk of re-injuries. Ekstrand found variations in re-injury rates between teams from different countries, with Danish teams showing an average 30% re-injury rate, Spain, Holland and England 19% and teams from France and Italy 11%. This could be due to the varying rehabilitation programmes employed by clubs in these countries, the styles of play adopted by each club or the different physical demands placed on players. Schmitt, Tyler & McHugh (2012) suggest a rehabilitation programme that includes eccentric hamstring exercises (such as the Nordic Hamstring exercise) to increase eccentric hamstring strength when in a lengthened state, may reduce risk of re-injury. The use of eccentric hamstring exercises as part of a rehabilitation protocol has gained support in recent years with several studies advocating their use once an individual has reached Phase Two of their rehabilitation programme (Lorenz & Reiman, 2011; Heiderscheit et al., 2010; Comfort, Green, & Matthews, 2009).

There is a wide body of research that suggest re-injuries are a major risk to players and cause them to have longer absences from the game than non-re-injuries. This highlights the important role that management and staff should play to help avoid re-injury, by employing clear, consistent and realistic rehabilitation programmes that have the player’s welfare as its focus. As outlined by Woods et al. (2002), injury prevention should place a heavy emphasis on effectively controlling players, with management having a major influence on the behaviour of players in the design of training programmes and the safety culture it promotes.

Exposure

The number of matches players participate in over a season is the subject of much debate within soccer, with many managers and coaches at the elite level of the game suggesting players are over exposed to matches season upon season. In an attempt to normalise injury rates, many research studies calculate the number of injuries per training or match exposures. This makes allowances for the increased injury risk faced by those players involved in higher levels of training and matches. Arnason et al. (2004) reported that players in their injured group had a significantly higher match exposure than those not injured (p=0.004). When considering training exposure however, uninjured players had a higher mean training exposure (65.3 hours) compared to injured players (62.5 hours). Match injuries were more common than training injuries (one injury per 1.2 matches and 19.3 training sessions). The lower training exposure for injured players could be explained by the fact that their injuries caused them to miss more training sessions, rather than the training sessions making non-injured players more injury resilient. Some caution should be taken when considering the training exposure data from this study due to the fact training exposure was recorded by the teams coaching staff and training attendance from five of the teams was somewhat incomplete. Training exposure was therefore estimated based on the team's schedule and feedback from the coaches.

Soderman et al. (2001), in their study of 146 players, reported that when exposure to soccer was analysed with other key variables such as age, height, and weight, it was found to significantly increase the risk of traumatic injury. These findings were recorded from a mean training exposure of 66.9 hours and match exposure of 32.9 hours per player per season. Ekstrand (2008) suggested that match exposure was not an issue in relation to injuries and performance for elite level teams. He found an average of 36 matches per player per season, which is significantly lower than the 60 to 70 matches played by the team during the same period. A player's absence through injury or non-selection would have skewed these average match figures, and as such the actual matches played for regular starters was likely to be higher than 36. This mean individual reduction in match exposure was similar across different countries, meaning clubs tended to deal with fixture congestion by increasing squad sizes rather than over exposing their players. Ekstrand also reported that injury risk was no higher in the final 10 weeks of the season in his study, suggesting that elite players in today's game can cope with the match schedules they are faced with.

Soderman et al. (2001) reported 10 traumatic injuries per 1000 player hours during matches and 1.3 injuries per 1000 player hours during training. Arnason et al. (2004) reported an injury incidence of 24.6 injuries per 1000 match hours and 2.1 injuries per 1000 training hours, and Ekstrand (2008) reported 24 to 30 injuries per 1000 match hours and 3 to 5 per 1000 training hours. This data clearly shows that injury risk during matches is significantly higher than in training despite the exposure to match hours being lower during a normal season. The intensity levels and increased contact during matches would seem to be the main cause of these differences. The training injury incidence rates remain relatively low compared with the exposure levels across a season and perhaps highlight the difficulty in replicating match like intensity levels during practice sessions. It should also be noted that the ability level of the players in Ekstrand's study was significantly higher than in the other two studies, and involved a different gender, and as such the speed and intensity levels faced by these players would be higher and increase injury risk.

Whilst the research suggests that there are certain risk factors that could predispose a player to injury, such as previous injury, it seems that there is no single risk factor that can explain the incidence of injuries in soccer. A multitude of factors conspire to increase injury risk across all levels. Other intrinsic biomechanical and physiological risk factors such as hamstring/quadriceps ratio, joint laxity, pelvic tilt, and foot pronation, have all been suggested as potential risk factors for injury, along with more extrinsic factors such as playing surface, warm up, weather conditions and many others. As the research into injury types, risk factors and prevention/intervention programmes continues to grow, practitioners are now better placed to manage their player's training programmes to actively help prevent injury and improve rehabilitation techniques. This improved knowledge and understanding should lead to reductions in the number of injuries and length of absence due to injury in the foreseeable future.

2.1.5 INJURY TRACKING

As identified in the previous pages, tracking injury rates and comparing results between studies can be a difficult process. Accuracy of injury reporting is critical, as any errors in this regard will affect the validity of any results published. As many surveillance studies rely on input from a third party when tracking injury, it is not always possible to determine the accuracy of the injury data collected. It is also important to apply appropriate tracking methods in areas such as injury recording, study length, sample sizes and injury definition, to allow for suitable analysis to be carried out on the results. To date there has been variation between most injury surveillance studies that have applied different collection methods, sample sizes and injury definitions, as highlighted in Table 2.2.

Table 2.2. Differences found in length of injury surveillance, definition of injury and collection of data amongst injury surveillance studies

Authors and Date	Sport	Length of Injury Surveillance	Sample Size	Definition of an Injury	Data Collection Method
Brummitt et al., 2013	Various Sports	One Season	193 athletes	Lower back or extremity injury that required athlete to be removed from activity or to miss next activity	Recorded by university athletic training staff
Chorba et al., 2010	Soccer, Volleyball and basketball	One Season	38 athletes	Injury required medical attention from club medical staff	Recorded by certified athletic trainers
Ekstrand et al., 2009	Soccer	Seven Seasons	2226 players	Player unable to participate in soccer activity	Recorded by club medical staff
Hawkins et al, 2001	Soccer	Two seasons	2376 players	Unable to participate in soccer activity 48 hours after injury occurred	Recorded by club medical staff
Kiesel et al., 2007	American Football	One Season	46 players	Placed on the injured reserve and time loss from playing of at least 3 weeks	Recorded by club medical staff
Newell et al., 2006	Gaelic Football	One Season	511 players	Unable to participate in soccer activity 48 hours after injury occurred	Recorded by club physio
Nilstad et al., 2014	Soccer	One Season	173 players	Unable to participate in soccer activity one day after injury occurred	Player recorded injury
Schnieders et al., 2009	Rugby Union	One Season	271 players	Player required medical attention during a match or resulted in player missing one training session or match	Recorded by club medical staff

Sample sizes from the studies in Table 2.2 vary dramatically and make it difficult to make comparisons between studies. Those studies with larger sample sizes, such as those produced by Ekstrand et al. (2009) and Woods et al. (2004), provide useful data from which clear trends can be seen and direct comparisons can be made. Unfortunately, such large sample sizes are uncommon in injury surveillance studies, due to the obvious difficulties in collecting accurate data from such large numbers and the manpower required to conduct such a study. When making comparisons between such studies, it is important to understand the differences that may exist between them and the effect these have on the results published. If these are fully understood, then it can be possible to draw certain conclusions from these studies.

The duration of surveillance can cause inaccuracies when comparing injury data across studies. The majority of studies in this area have undertaken surveillance for one season, but as different sports have varying season lengths, the actual number of days these studies lasted varies significantly. As injury risk and patterns vary across a season, it is recommended that injury surveillance studies last for an entire season or over several seasons (Hägglund et al., 2005). It appears the longer the study can last the more useful the data will be, simply due to the likely increase in the number of injuries reported over a longer period, which will effectively provide a researcher with a larger sample size and dataset.

Injury definition is perhaps the biggest factor that causes discrepancies between data from different studies. As Table 2.2 shows, there have been several definitions used in injury surveillance studies due to the lack of any standard definition being in existence. This makes it very difficult to accurately compare studies when considering the number and types of injuries reported, as the definition used in each may vary. The most common injury definition is the time loss definition (Woods et al., 2004; Ekstrand et al., 2009; Nilstad et al., 2014) which means that an injury is defined as one that results in a player being unable to fully participate in training sessions or matches. However, other studies have used a medical assistance definition (Junge & Dvorak, 2013) that means an injury was defined as any physical complaint incurred during a match that received medical attention from the team physician. A limitation of the time loss method is its inability to track ongoing overuse injuries. These type of injuries may cause athletes pain and force them to adapt their training routines, but do not always result in them ceasing participation in their chosen sport (Clarsen, Myklebust, & Bahr, 2013). Typically athletes will defer time loss and often wait until off-season periods to take the rest required to treat overuse injuries, which are rarely covered in injury surveillance studies (Clarsen et al., 2013). From an injury reporting point of view the time loss definition is most commonly used, partly because it allows the researcher to calculate injury severity, and the medical complaint definition is often only reported during tournament periods. Clearly the definition employed in a study will result in differences in the number of injuries reported, which leads to possible discrepancies when comparing data from studies that have used different definitions.

Lastly, the manner in which injuries were reported differs across research. Many studies relied on the medical staff from each club to monitor and record injury data (Walden et al., 2005; Ekstrand, 2008). Although this leaves the reporting procedure outside of the control of the researcher, the nature of such studies means this is the most reliable and accurate method to use to report injuries. Other studies used a self-reporting method, which relied on the individual players to report injuries (Nilsted et al., 2014). This method could lead to many injuries being unreported as players are left to their own discretion in terms of what they consider an injury to be.

Due to these issues, there has been an attempt by both FIFA and UEFA to standardize the injury surveillance process, to help make comparisons across studies easier and more relevant. The UEFA Medical Committee held discussions in 1999 and 2000 to discuss optimal study designs and reporting methodology. This led to the development of a questionnaire and two pilot studies being run to test study design and data collection methods. By May 2001, the proposed methodological design, definitions and reporting forms were approved after minor revision (Hagglund et al., 2005). As outlined in the review by Hagglund et al. (2005) the UEFA model makes several recommendations relating to practical guidelines for epidemiological studies of football injuries, outlined in Table 2.3.

Table 2.3. Summary of main recommendations included with UEFA Injury model (Hagglund et al., 2005)

UEFA INJURY MODEL RECOMMENDATIONS	
Study Design	Studies should have a prospective cohort design
Exposure Factor	Exposure factor, that is the time during which the participant is at risk of injury, should be measured
Study Period	As a minimum, studies should last for an entire season, including pre-season, and ideally last over many seasons.
Data Collection Forms	It is recommended that three forms are used for the collection of data in a study of injury risk in sports:
1. Baseline Form	To record anthropometric data and previous medical history
2. Exposure Registration Form	To record individual attendance and exposure over each season, and covers both training sessions and matches
3. Injury Form	To record information relating to injuries that occur including date, type, location etc.
Study Manual	An instruction manual that describes the various definitions used in the study in detail
Contact at Club	Each club should nominate a contact to collect all data, and ideally this should be a member of the medical team
Injury Definition	A time loss definition of injury should be used, meaning that a recordable injury is one that causes absence from the next training session or match soccer and occurred during a scheduled training session or match.
Injury Severity	It was agreed that injury severity would be classified based on the number of days of absence from real participation as follows:
1. Slight	1 to 3 days absence
2. Minor	4 to 7 days absence
3. Moderate	8 to 28 days absence
4. Severe	More than 28 days absence

In addition to these recommendations from UEFA, further work was completed by an Injury Consensus Group put together by FIFA (Fuller et al., 2006). The group produced a statement which provided key definitions relating to injuries that could be used by surveillance studies to allow accurate comparison to be made across research. The key definitions suggested were outlined by Fuller et al. (2006) as shown in Table 2.4.

Table 2.4. Summary of key definitions outlined by FIFA injury consensus group (Fuller et al., 2006)

FIFA Injury Consensus Group Key Definitions	
Injury	Any physical complaint sustained by a player that results from a soccer match or training, irrespective of the need for medical attention or time loss from soccer activities.
Recurrent Injury	Defined as an injury of the same type and at the same site as an index injury and which occurs after a player's return to full participation from the index injury.
Injury Severity	Number of days that have elapsed from the date of injury to the date of the player's return to full participation in team training and availability for match selection.
Match Exposure	Defined as play between teams from different clubs.
Training Exposure	Team and/or individual physical activities under the control or guidance of the team's coaching or fitness staff
Injury Classification	Injuries should be classified by location, type, body side and mechanism of injury and whether the injury was a recurrence.
Study Population	The study population should normally consist of more than one team of players and the study should last for a minimum period of one season (including pre-season), 12 months or the duration of a tournament.
Data Collection	Standardised forms should be used to collect data from the sample population. Should include; Baseline information form, Injury form and a Match/Training exposure form.

These two models provide clear guidance regarding the implementation and operation of injury surveillance studies in soccer and the best practice to employ to improve the accuracy of the data collected across research studies. However, slight variations exist between each model, and could lead to minor differences in the way injuries are defined and the injury data collected. This makes it difficult to implement all the recommendations suggested in both models when conducting a study, and would potentially lead to one model being used over the other when undertaking an injury surveillance study.

2.2 MOVEMENT SCREENING

Movement evaluations and pre-participation screens have been widely adopted by scientists and practitioners due to the fact that links have been made between individual's movement behaviours and their risk of injury (Frost, Beach, Callaghan, & McGill, 2012). Such procedures are common in a sporting context and are widely used during pre-season periods to establish dysfunction, compensations and asymmetries within the body that could impact on performance. Typically, the traditional medical model emphasizes identification of an anatomical source of pain via assessment of a tissue and/or task specific to the impaired joint (Glaws, Juneau, Becker, Di Stasi, &

Hewett, 2014). Such a model may not be optimal for the management of musculoskeletal pathology, as it does not account for 'regional interdependence' or the concept that adjacent anatomical regions can contribute to or be the source of a patient's primary complaint (Wainner, Whitman, Cleland, & Flynn, 2007). In addition, many such screening methods can involve expensive and complex procedures to achieve accurate measurements, making them prohibitive in applied settings.

In order to attempt to screen for factors relating to injury, professionals have begun to assess function using foundational movement patterns that require co-ordinated utility of multiple joints and their movements (Gulgin & Hoogenboom, 2014). Screening and assessment tools that incorporate whole body functional movements may uncover important underlying impairments that allow for the development and implementation of targeted interventions (Glaws et al., 2014). With a whole body, multi-joint, analysis of functional movement becoming more prevalent in the field, there has been a steady increase in the amount of research relating to movement screening.

2.2.1 SCREENING TECHNIQUES

Many screening methods and tests are able to detect risk of a specific injury or weakness. As highlighted in earlier sections of this chapter, the thigh, ankle and knee remain the most common injury locations within soccer (Ekstrand et al., 2009). Therefore, for the purposes of this programme of research, it was important to understand which screens/tests were able to reliably predict injuries within these locations, and to provide a critique of such screens, to provide the reader with a clear understanding of their effectiveness in applied settings.

The Ankle

It has been reported that ankle injuries can make up 17% of all injuries reported in soccer (Hawkins et al., 2001). As soccer is a sport that involves unilateral balance and co-ordination, a screen that can assess an individual's unilateral balance could provide a useful indication of an individual's risk of receiving an ankle injury. The Star Excursion Balance Test (SEBT), is an inexpensive, quick method of measuring balance, with good reliability (Plisky et al. 2006). The SEBT involves an individual balancing on one leg, whilst they attempt to reach in eight different directions with the opposite leg. The maximum reach distance achieved in each direction is recorded, and a total composite score for all reach distances is calculated and compared to limb length. Olmstead et al. (2002) found players with chronic ankle instability had significantly decreased reach distances compared to both healthy individuals, and their opposite limb. Due to the time and set up processes involved when using eight reach distances, a modified version of the SEBT was developed using only three reach distances. Hertal et al. (2006) reported that the posteromedial reach direction identified individuals with chronic ankle instability compared to healthy controls. Further research using the modified SEBT also found an increase in lower extremity injury risk in basketball players with either a decreased normalised right composite reach score or an anterior reach right/left distance difference of greater than 4cm (Plisky et al., 2006).

In recent years, the modified SEBT has been further developed by adding a testing kit and is now referred to as the Y-Balance test (YBT). The YBT involves strength, flexibility, neuromuscular control, stability, range of movement, balance and proprioception (Gonnell et al., 2015). Research has suggested that performance on the YBT can be linked to injury risk with Gonnell et al. (2015) reporting that players with a right/left difference in the posteromedial reach direction of 4 cm or greater were 3.86 times more likely to sustain a lower extremity injury. Additional research by Smith et al. (2015), found that an asymmetry of 4 cm in the anterior reach direction was the optimal cut-off point for predicting injury, whereby subjects with an asymmetry of 4 cm or greater, were at an increased risk of receiving a lower extremity injury. These authors also suggested that it is not possible to apply one composite score risk cut-off across multiple sports. This implies that further research is required before the YBT can be used accurately across multiple sports to predict an athletes risk of an ankle or lower extremity injury. More recent research studying the YBT and lower extremity injury risk, found no significant relationship between occurrence of a lost time lower extremity injury and performance on the YBT (Walbright et al., 2017).

Whilst it appears that performance on the SEBT and /or YBT has some link with injury risk, there are differences in regard to the cut-off points and reach directions reported for both total composite score and asymmetry between right/left sides. Although it has been suggested that such cut-off points vary across different sports and genders, this appears somewhat difficult to substantiate. Further work in this area is warranted to fully understand if the mechanisms that lead to ankle and lower extremity injuries change across sports or genders.

The Knee

Knee injuries can account for 18% of all injuries reported in soccer (Ekstrand et al., 2009). It is a particular issue for female players, with deceleration, lateral pivoting and landing tasks reported as significant risk factors for Anterior Cruciate Ligament (ACL) injuries (Hewett et al., 2005). Knee valgus lower limb alignment is often seen during non-contact (ACL) injuries when landing or jumping (Noyes et al., 2005). Therefore, screening tests to identify knee injury risk tend to focus on landing or jumping activities, with the drop jump and tuck jump tests, the most commonly employed techniques to date.

The drop jump test involves a participant jumping from a box 30 cm high, landing bilaterally in front of the box, before performing a maximum vertical jump. Variations of this protocol and the variables measured have been utilised in studies, when using the drop jump to assess landing mechanics. Noyes (2005) used the protocol outlined in the previous sentence and assessed the absolute separation distance between the left and right hip, and the normalised separation distances for the knees and ankles. The author used these measures as they believed it would provide an indication of the participant’s ability to control lower limb axial alignment in the coronal plane and reported excellent levels of test-retest and within-test trial reliability for all measures. Although the drop jump protocol used in this study provided an indication of the difference in landing mechanics between males and females, the authors accept that the test could not be used as a risk indicator for knee ligament injury as the single camera analysis only depicts hip, knee and ankle positions in a single plane. Hewett et al. (2005) used a similar protocol to Noyes (2005) in their study involving 205 female adolescent soccer, basketball and volleyball players. However, one major difference compared to the Noyes (2005) study was the inclusion of 3D biomechanical analysis using 25 reflective markers and a multi camera motion capture system, along with 2 force platforms to gather reaction force data. This allowed precise joint kinematic and kinetic analysis to be conducted such as measuring hip adduction, knee abduction and ankle eversion. The authors reported that injured subjects demonstrated significant increases in dynamic lower extremity valgus and knee abduction loading before sustaining injuries compared to uninjured controls.

Further research in this area incorporated the drop jump protocol described in the previous paragraph, along with a single leg landing protocol, to assess knee valgus alignment in 50 male and 50 female subjects (Herrington & Munro, 2010). The single leg landing protocol involved participants stepping from a 30 cm box and landing on their opposite leg, holding the position upon landing, to simulate typical landings encountered during athletic performance. The authors used 2D frontal projection plane video analysis to assess knee valgus angle during test performance. The results indicated no difference in knee valgus angle between genders for both the bilateral drop jump and unilateral step landing tests. The authors suggest a more effective way of establishing injury risk is to establish population norms and then to assess if individuals outside these norms are at a higher risk of injury. For the drop jump test, the authors suggest that knee valgus angle should be symmetrical and in the range of 7 to 13° for females and 3 to 8° for males, and for the unilateral landing task knee valgus angle should be symmetrical and in the range of 5 to 12° for females and 1 to 9° degrees for males.

Research to date relating to bilateral drop jump tests would suggest that the test can be used to provide an indication of knee injury risk, predominantly in female athletes, particularly when using more sophisticated analysis techniques such as 3D motion capture and force platforms. Increased knee valgus and knee abduction angles during landing have been reported as an injury risk factor in female participants. However, from the current literature in this area, it is difficult to effectively establish a definitive range of measures that provide a clear indication of injury risk, from simply using the drop jump test.

An additional tool that is often used to assess risk of ACL injuries is the tuck jump assessment. It has been found that the risk of ACL injuries can be reduced through adopting neuromuscular training that incorporate high intensity plyometrics (Myer, Ford, & Hewett, 2008). These authors suggest the tuck jump exercise may be a useful tool for the identification of lower extremity landing technique flaws during plyometric activity. Participants initiate the jump with a slight crouch downward, with arms behind them. They then swing their arms forward, whilst jumping straight up at the same time, pulling their knees as high as possible. At the highest point, the athletes are in the air with thighs parallel to the ground. Upon landing the athlete repeats the same process immediately (Myer et al., 2008). These authors developed a 10-point scoring criteria to be used during 10 seconds of tuck jump performance, with the assessment completed manually by a rater in real time, or by using video footage captured in the frontal and sagittal plane. Flaws on six or more of the scoring criteria would indicate a higher risk of injury according to the authors, based on empirical evidence gathered within their lab. They also reported a high inter-rater reliability for the tuck jump, suggesting the tuck jump assessment could be used reliably by a single rater to reassess performance.

Additional research to assess the intra and inter-tester reliability of the tuck jump assessment has been conducted (Dudley et al., 2013; Herrington, Myer, & Munro, 2013). These studies incorporated the testing and scoring protocols developed by Myers et al (2008) as discussed in the previous paragraph. Herrington et al (2013) reported that the tuck jump assessment showed very good to excellent intra and inter-tester reliability when two raters assessed 10 subjects using video footage collected from two cameras positioned in the frontal and sagittal plane. Dudley et al (2013) utilised the same tuck jump assessment protocol to assess 108 recreationally active students. From these performances, videos of 40 participants were randomly selected in order to test the inter and intra-rater reliability of five raters when using the tuck jump assessment. The results indicated poor to moderate intra-rater reliability for three of the raters, and poor inter-rater reliability between all five raters. These authors suggest the larger sample size used in their study compared to that of Herrington et al (2013) and the limited tuck jump assessment experience of their raters could be factors that affected the lower reliability results reported.

To date there is limited research relating to the tuck jump assessment and its ability to accurately predict injury risk across a range of samples. Whilst some studies have suggested it can be used reliability by single or a group of raters (Myers et al., 2008; Herrignton et al., 2013), there is as yet, no research that has clearly identified its ability to predict ACL injuries. Although the tuck jump assessment could potentially offer a time efficient and simple solution, further research is required to fully understand its injury prediction capabilities.

The Hamstrings

Hamstring injuries remain the most common injury location within soccer, accounting for 23% of all injuries reported (Ekstrand et al., 2009; Hawkins et al., 2001). It has been reported that the majority of hamstring strains occur during high speed running, which result in the sudden onset of posterior thigh pain (Heiderscheit et al., 2010). The risk factors associated with hamstring injuries are multi factorial in nature and include quadriceps peak torque, older age and previous hamstring injury (Freckleton, Cook, &

Pizzari, 2014). These authors also suggest limb dominance, playing position. ethnicity, ankle dorsiflexion, range of movement and previous knee injury have limited evidence to support their association with hamstring injury risk. Such varied risk factors make the hamstring a difficult injury to screen for, and this has resulted in a wide range of tests being used in attempt to screen for potential hamstring injury risk.

The sit and reach test, active straight leg raise and passive straight leg raise tests, are often used to assess hamstring flexibility. The sit and reach test is designed to measure hamstring and lower back flexibility, and it is often believed that maintaining hamstring and lower back flexibility may prevent acute and chronic musculoskeletal injuries (Baltaci, Un, Tunay, Besler, & Gerçeker, 2003). The test involves participants placing the sole of one foot against the end of a standard box, with the leg fully extended. Placing one hand on top of the other with palms down, the participant reaches forward as far as possible along the measuring scale without bending the extended knee or moving the sole of the foot. Using this protocol, Baltaci et al. (2013) reported that the sit and reach test was highly related to hamstring flexibility, when tested on 102 female university students. A meta-analysis conducted by Mayorga-Vega et al. (2015) that considered 34 studies, reported that the sit and reach test had a moderate mean criterion related validity for estimating hamstring extensibility, further suggesting that the test is a valid tool to use in order to assess hamstring flexibility. Additional flexibility tests have also been incorporated into screening programmes to assess hamstring injury risk. The active straight leg raise involves the participants lying in a supine position, with arms by their side and toes facing upwards. With one leg remaining still on the floor, and no additional head or trunk movement, the participant raises the other leg to maximise hip flexion whilst maintaining knee extension. The procedure is then repeated on the opposite leg. The greater hip flexion achieved whilst maintaining knee extension, the better the level of hamstring flexibility and hip mobility within the participant. The passive straight leg raise follows the same procedure, but involves a tester assisting the participant with their leg raise, so they remain passive during performance of the test, to remove any motor control issues the participant may have when performing the raise actively. Sorani & Rathod (2016) assessed the passive and active straight leg raise tests, alongside passive and active knee extension tests, in their study involving 100 participants, to determine hamstring range across all four tests. They reported significant correlation between the tests, suggesting that the active and passive straight leg raise tests are valid measures of hamstring flexibility.

Whilst it appears that the sit and reach and straight leg raise tests provide a good indication of hamstring flexibility or extensibility, it is still unclear how/if these measures are linked to hamstring injury. As previously outlined, hamstring injury is multi-factorial and therefore it is unlikely that a test that only measures hamstring flexibility could detect hamstring injury risk with any great accuracy. To date there is a lack of research that has considered the ability of the sit and reach test and straight leg raise tests to predict hamstring injury.

Strength is also perceived to be a risk factor for hamstring injury, despite evidence being inconclusive, as it is possible that a weak muscle may fatigue earlier during activity and therefore have to work harder than its physiological capacity to maintain performance during high intensity activity (Freckleton et al., 2014). These authors used the single leg bridge test, a clinical test for the hamstring where the hip and knee are at functional angles, to assess risk of hamstring injury in 482 amateur and semi-elite Australian Rules Football players. Previous research by Hallet (2010) reported high levels of inter-tester (ICC=0.77-0.89) and intra-tester (ICC=0.89-0.91) using a single leg bridge protocol that involved participants lying down with one heel on a 60 cm high box and with the knee in approximately 20° of flexion. Participants crossed their arms over their chest and were then instructed to push down through the heel of the test leg to lift their bottom off the ground, whilst keeping the non-working leg as still as possible and repeat for as many repetitions until they reached their failure point. This protocol was utilised in the study by Freckleton et al. (2014) and it was reported that players sustaining an injury on their right leg had a significantly lower mean single leg bridge score on this leg compared to non-injured players. Such a relationship was not found for the left leg. Further univariate analysis of the results from this study suggested that the single leg bridge may be a useful screening tool, when assessed alongside age and previous injury, although it was not possible to isolate the contribution of the single leg bridge within this multi risk profile.

Whilst the single leg bridge test has shown some promise as a predictor of hamstring injury, further research is required, that follows the strict protocol described above, before it is possible to understand if the test can accurately predict risk of hamstring injury, either as a standalone measure or when used as part of a univariate analysis involving several injury risk variables.

Whilst the previous paragraphs have highlighted a number of screens that can be undertaken to specially focus on ankle, knee and hamstring injuries, the results from the research to date highlight that further work is required in order to accurately use these screens in applied settings on a regular basis. Research has identified that tests such as the Y Balance and Drop Jump are able to be used to identify injury risk of the ankle and knees respectively. However, the results do not currently provide a definitive range or cut-off point that can be used across all samples to accurately predict injury risk. An asymmetry in reach difference between the left and right leg on the Y Balance, and the knee valgus angle or knee separation distance on the drop jump tests, appear to be the key measures that could be used to identify injury risk, if standardised cut-off points or range of measurements could be established, that apply to a wide range of age groups, sports and both genders.

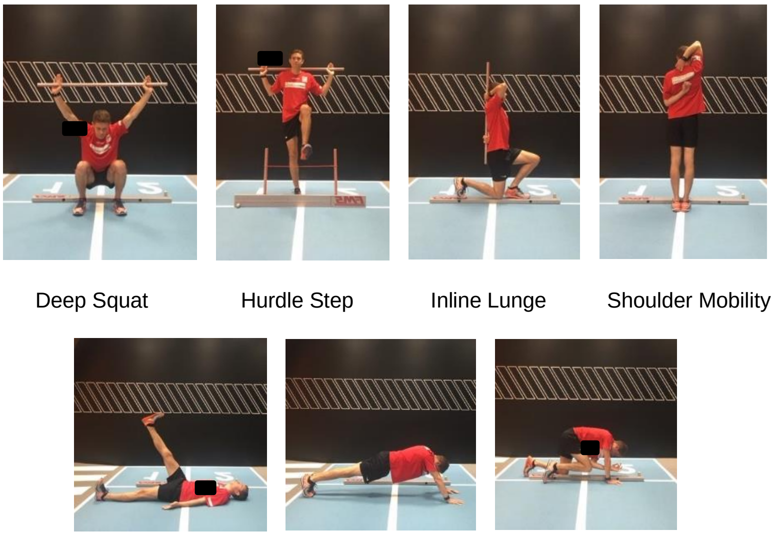
Additionally, the tests discussed in the previous paragraphs typically focus on one injury location, which would therefore require a range of tests to be conducted in order to accurately screen for all type of injury risk. From a practical perspective this can become time consuming and lead to an increase in the possibility of inter-rater and intra-rater error due to multiple tests being conducted. However, a multi-test approach is now common place within elite sport as medical and sport science staff strive to reduce injuries across a squad of players. These multi-test protocols often assess movement and motor control capability, rather than specifically attempting to assess injury risk related to a particular injury type or location. Frohm, Heijne, Kowalski, Svensson, & Myklebust (2012) developed a nine-test screening battery in their study involving 18 male elite soccer players, which consisted of functional and complex movement exercises picked from different test batteries that had been tested and retested over a 10 year period. The nine tests included screens designed to assess mobility, stability and flexibility in both bilateral and unilateral positions. Using eight physiotherapists, this study assessed the inter-rater and intra-rater reliability across two testing sessions, seven days apart, and reported good levels of reliability between the raters and between sessions. Whilst this study shows that the nine-test screening battery had good levels of reliability when used in this sample, further research with larger samples, that have more varied performance levels, are needed before the battery can be deemed to be reliable. Additionally, as the authors outline, this study did not assess the nine-test battery’s ability to predict injury, and as such, it is not possible to state if using this protocol could accurately identify injury risk.

A more recent study developed The Athletic Ability Assessment (AAA) to be used as an assessment methodology to be utilised as athletes travel along their sporting pathway and require greater levels of movement competency (McKeown, Taylor-McKeown, Woods, & Ball, 2014). Each of the seven movements within the AAA assesses a range of functional movement qualities such as trunk stability, hip, knee, ankle alignment, squat or lunge ability and the ability to jump and land correctly. Using 17 female football players as participants, McKeown et al. (2014) assessed the inter-tester and intra-tester reliability of the primary researcher using real time scoring and on two separate occasions, video scoring, compared to video scoring by five additional raters. They reported high levels of agreement between scorers for total scores given for all seven tests. This level of agreement reduced when considering individual test scores. The AAA is a novel approach to the assessment of movement qualities related to sporting performance, but additional research is required, on larger samples, from different sports and with a varied performance level, before it can be deemed to be a reliable and valid protocol to use in applied settings. Additionally, the AAA has not been used to assess injury risk in any study to date, and therefore it is not possible to suggest that the protocol can be used to accurately predict injury.

Multi-test approaches are becoming more common in applied settings, with many designed to assess various movement capabilities that underpin sporting performance. To date, many of these protocols do not have sufficient research to support their use in applied settings, either from a reliability/validity or an injury prediction standpoint. However, one particular seven test movement screen, the Functional Movement Screen (FMS) has received significant attention over the past 10 years, with several research studies undertaken to assess its reliability as a screen, its ability to predict injury and also its ability to predict performance. The following section discusses this movement screen in more detail.

2.2.2 FUNCTIONAL MOVEMENT SCREEN (FMS)

One specific seven subtest Functional Movement Screen (trademarked as the FMS) has gained particular popularity (Cook, Burton, & Hoogenboom, 2006), with the screen regularly being used in a sporting context, by sports science staff, as a pre-participation screen, and by medical and health professionals as part of a return to play protocol to assess an individual’s recovery from injury. This screen is designed to help health professionals identify any dysfunction, asymmetry or weakness that may exist within an individual’s movement patterns, helping to direct future assessment or treatment. Each of the seven subtests are scored between zero and three and assess a person’s mobility, stability and flexibility during multi-joint movements. The seven subtests are the Deep Squat (DS), Hurdle Step (HS), Inline Lunge (IL), Shoulder Mobility (SM), Active Straight Leg Raise (ASLR), Trunk Stability Push Up (TSPU) and Rotary Stability (RS).



Deep Squat Hurdle Step Inline Lunge Shoulder Mobility

Active Straight Leg Raise Trunk Stability Rotary Stability

Figure 2.1. The seven FMS subtests

Table 2.5. The scoring criteria for the seven FMS subtests

Test	Score of 3	Score of 2	Score of 1
Deep Squat	1. Femur below horizontal 2. Knees aligned over feet 3. Dowel aligned over feet 4. Upper torso parallel with tibia	As per 3 but with heels placed on a 6x2 board	Unable to perform movement without compensation
Hurdle Step	1. Hips, knees and ankles remain aligned in sagittal plane 2. Minimal to no movement in lumbar spine 3. Dowel and hurdle remain parallel	1. Alignment lost between hips, knees and ankles 2. Movement in lumbar spine 3. Dowel and hurdle not parallel	1. Contact between foot and hurdle 2. Loss of balance
Inline Lunge	1. Dowel contacts remain with L Spine extension 2. No torso movement 3. Dowel and feet remain in sagittal plane 4. Knee touches board behind heel of front foot	1. Dowel contacts do not remain with L Spine extension 2. Movement in torso 3. Dowel and feet not in sagittal plane 4. Knee does not touch board behind front foot	Loss of balance is noted
Shoulder Mobility	Fists are within one hands length	Fists are within one and a half hands length	Fists are not within one and a half's hand length
Active Straight Leg Raise	Ankle/dowel resides between mid-thigh and ASIS	Ankle/dowel reside between mid-thigh and mid patella/joint line	Ankle/dowel resides below patella/joint line
Trunk Stability Push Up	1. Males perform 1 rep with thumbs aligned with top of forehead 2. Females perform 1 rep with thumbs aligned with chin	1. Males perform 1 rep with thumbs aligned with chin 2. Females perform 1 rep with thumbs aligned with clavicle	1. Males unable to perform 1 rep with thumbs aligned with chin 2. Females unable to perform 1 rep with thumbs aligned with clavicle
Rotary Stability	1. Performs 1 correct unilateral rep whilst keeping spine parallel to board 2. Knee and elbow touch in line over board	1. Performs 1 correct diagonal rep whilst keeping spine parallel to board 2. Knee and elbow touch in line over board	Inability to perform diagonal reps

More than 400 citations relating to the FMS were identified between 2004 and 2013, from 33 publications (Kraus, Schultz, Taylor, & Doyscher, 2014) and many more research articles have been published since 2013 relating to this area. These publications primarily focus on three key areas; FMS and Injury, FMS and performance, and the reliability of the FMS. This level of research makes the FMS a popular choice compared to some of the screens/tests discussed earlier in this chapter. One of the main benefits of using the FMS is the standardised scoring criteria applied to each of the seven subtests. This provides a practitioner with a simple to follow procedure when scoring each performance on every subtest, helping to increase reliability and reduce scoring errors. In addition, the FMS uses an easy to set up and assemble testing kit for all seven subtests, which allows the screen to be completed efficiently and without the need to apply body markers or set up cameras to record performance. From a practical perspective, the FMS offers practitioners the opportunity to follow a standardised protocol, supported by extensive research, that provides an indication of both injury risk and performance levels, that other screening tools cannot currently offer, due to a lack of research. Finally, the FMS assesses non-sports specific, multi-joint movement, specifically mobility, stability and flexibility, to understand any weaknesses or asymmetries that may exist. This provides practitioners with a clear indication of where to focus their intervention programmes in an attempt to improve an individual's fundamental movement, which underpins all sporting actions. Compared to tests that offer single joint assessments focussed on one particular joint or injury type, the FMS considers the impact of regional interdependence and the affect movement compensations could have on each participants ability to move efficiently and free of pain.

2.2.3 LINKS BETWEEN FMS SCORE AND INJURY RISK

As previously discussed, reducing injury rates and improving performance are obvious objectives for any practitioner working with individuals involved in physical activity or highly physical occupations. Research has focussed on the ability of the FMS to predict injury and performance levels across various populations. Table 2.6 highlights research within this area, and the differences that exist in the methods employed by each study.

Table 2.6. Research analysing links between FMS scores and injury

Authors and date	Population	Length of injury surveillance	Sample size	Definition of an injury	Exposure levels reported	Link with injury reported	Cut-off score applied	Odds ratio reported	Injury type/severity
Chorba et al., 2010	Soccer, Volleyball and basketball	One Season	38	Injury required medical attention from club medical staff	No	Yes	14	3.9	All musculoskeletal injuries
Duke at al., 2017	Rugby	One Season	73	Player unable to take a full part in future training or match play	No	Yes	14	7.7	All injuries
Garrison et al., 2015	NCAA Division 1 and club sports	One Season	160	Musculoskeletal pain during athletic participation and adjusted training for at least 24 hours	No	Yes	14	5.6	All musculoskeletal injuries
Kiesel et al., 2007	American Football	One Season	46	Placed on the injured reserve and time loss from playing of at least 3 weeks	No	Yes	14	11.7	Injuries lasting 21 days or more
Letafatkar et al., 2014	Physically active students	One Season	100	Any lower extremity injury that resulted in no participation for one or more exposures	No	Yes	17	4.7	All lower extremity Injuries
O'Connor et al., 2010	Marine Officers	38 to 68 days	874	Subject sustained physical damage to the body and sought medical care one or more times	No	Yes	14	1.5*	All injuries
Peate et al., 2007	Firefighters	Retrospective	433	Not defined	No	Yes	16	1.68	All injuries
Appel, 2012	Track and Field athletes	One season	57	Inability to participate in practice or competition for at least 4 days	Yes	No	13	0.48	All injuries
Bardenett et al., 2015	High School Athletes	One Season	167	Any musculoskeletal injury that resulted in one or more future exposures being missed	No	No	14	0.91**	All musculoskeletal injuries
Chalmers et al., 2016	Australian Football	One Season	237	A trauma or medical condition which caused a player to miss a competitive match	Yes	No	14	1.1	All injuries

*Relative Risk; **Positive Likelihood Ratio

Kiesel, Plisky, & Voight (2007) reported that individuals scoring 14 or less out of 21 on the FMS were over 11 times more likely to receive an injury in their study of 46 NFL football players. This study was the first of its kind to report a link between FMS and injury. No other study to date has reported an odds ratio as high as that seen in this paper. These results are questionable due to the limitations that exist within the methodology employed by the authors. This study was completed on a relatively low number of participants (n=46) over one season, and only included injuries that resulted in an absence of 21 days or longer, discounting any minor, slight and some moderate length injuries. In addition, no consideration was given to injury incidence compared to player exposure levels, and no reference was made regarding the differences in FMS scores and injury occurrence for players in different playing positions. These issues will be discussed in more detail later in this section.

Similar findings were reported by Chorba et al. (2010), who suggested a score of 14 or less led to a 4-fold increase in injury risk in their smaller study of just 38 female collegiate athletes. More recent research by Garrison, Westrick, Johnson, & Benenson (2015) also reported a link between a cut-off score of 14 and injury risk, in their study on 160 collegiate athletes. These studies appear to suggest that total FMS composite scores can be used to predict an individual's risk of injury and as such by improving an individual's FMS score, their risk of injury would be reduced. Moving away from athletic populations, studies have also identified a link between FMS and injury in physically demanding jobs such as firefighting (Peate, Bates, Lunda, Francis, & Bellamy, 2007) and marine officers (O'Connor, Deuster, Davis, Pappas, & Knapik, 2011). A study involving 433 firefighters being tested on the FMS suggested that the odds of scoring 14 or less was 1.68 greater for firefighters with a history of any injury (Peate et al., 2007). They also reported a decrease of 3.44 in the total FMS composite score for individuals with a history of injury. O'Connor et al. (2011) in a study of 874 marine officers reported that individuals scoring 14 or less on the FMS tests were 1.5 times more likely to receive any injury. This relationship was not apparent for overuse injuries.

Additional research in this area has provided contrasting results to those studies outlined in previous paragraphs, with Appel (2012) reporting no link between an individual scoring 14 or less and receiving an injury in their study of 57 collegiate track and field athletes. They reported no significant difference between mean total FMS composite scores for injured and non-injured players, and found a cut-off score of 18 or less as being more sensitive to predicting injury than a cut-off score of 14, but as being low in specificity. Bardenett et al. (2015) also found there were no statistically significant associations between total FMS composite scores and injury status, in a large study involving 167 high school athletes. The authors also found that injured athletes had significantly higher scores on the inline lunge subtest compared to non-injured athletes, with the reverse being found for the shoulder mobility subtest. More recent research by Chalmers et al. (2016) also highlighted the lack of association between total FMS composite score and injury risk. Their study on 237 elite junior Australian Rules football players found that a cut-off score of 14 or less was not linked to prospective injury risk, although players with at least one asymmetry experienced a significant moderate increase in sustaining an injury compared to players with no asymmetry.

The literature discussed in the last three paragraphs would suggest that FMS has some links with injury, and that a cut-off score of 14 could be a useful tool in helping to predict injury. However, caution should be applied when considering some of the research studies outlined in Table 2.6 due to limitations that exist within them. These include the injury definition used, comparison of injury type, within sample differences, lack of reference to exposure levels and the exclusion criteria applied.

Injury definition

As Table 2.6 shows, the definition of injury used within each study was different. Thus, although seven of the studies reported links between FMS and injury, it is very difficult to establish exactly how strong these links are. Perhaps the most prominent study in this area is that of Kiesel et al. (2007), which only considered injuries of 21 days or more. This means that only some moderate and all severe injuries were included in their study and all minor, slight and some moderate injuries were completely discounted. This compares to studies by Peate et al. (2007) that gave no injury definition, and O'Connor et al. (2011) and Chorba et al. (2010) that included any injury that required medical attention. More recent studies also provide conflicting definitions, with Garrisson et al. (2015) using a 24-hour absence from sports participation as an indication of injury, and Duke et al. (2017) using a player's inability to take part in future training or match play, with no time period provided, as their definition. It is clear the definition of injury must be carefully considered when attempting to interpret the results of previous FMS and injury research. Any future research in this area needs to carefully consider injury definition as well as making allowance for different injury types and durations when completing their analysis.

Injury type

The studies outlined in Table 2.6 all provided varying degrees of analysis of the type of injuries and their links with FMS scores. Kiesel et al. (2007) and Appel (2012) made no reference to injury type in their studies, Peate et al. (2007) referred to injury location but not type, and Chorba et al. (2010) looked at the relationship between anterior cruciate ligament injuries and FMS scores. Bardenett et al. (2015) considered only musculoskeletal injuries but did not narrow this definition to include a comparison between contact injuries versus non-contact injuries. Letafatkar, Hadadnezhad, & Shojaedin (2014) was the only study to consider the effect of contact versus non-contact injuries and their links with FMS score. They reported no significant difference in FMS score between individuals with a non-contact injury and those with a contact injury (p=0.22). The authors did report a significant difference in FMS score between individuals with a non-contact injury and those who did not receive an injury (p=0.03), but this was less significant than the difference in FMS score between individuals with a contact injury and those who did not receive an injury (p=0.01). Given that the FMS assesses fundamental movement patterns, it would be fair to assume that FMS scores should be more closely linked with non-contact injuries, and understanding the ability of the FMS to predict these would provide further insight into the link between FMS scores and injury risk.

Within sample differences

Limited allowance has been made by any study relating to differences that may exist within the sample population. For example, Kiesel et al. (2007) studied 46 NFL players but made no reference, or completed no analysis, relating to playing position and FMS scores. Significant differences exist in variables such as body composition, physical attributes (Schmidt, 1999), and FMS scores (Kiesel, Plisky, & Butler, 2011) between playing positions in American football. By pooling all players in the analysis together, Kiesel et al. (2007) may have created links between the FMS and injury risk that were actually driven by playing position. Interestingly the one study that did consider inter sample differences (Appel, 2012) reported no links between FMS scores and injury and could not establish an accurate cut-off score for their sample.

Exposure levels

As discussed previously in section 2.1.4, The UEFA injury model (Hagglund et al., 2005) recommends that when completing an injury surveillance study, a key factor to understanding injury occurrence rates is the level of exposure of the sample population. Specifically, it is important to establish the number of injuries per player per exposures, be it training or match/competition minutes. Only two of the studies highlighted in Table 2.6 considered the effect of exposure levels on injury. Appel (2012) considered exposure levels but completed no analysis or comparison of these in relation to FMS scores and injury occurrence. Chalmers et al. (2016) tracked player competition involvement in their study and used this information to assess how additional exposure affected the proportion of players not sustaining an injury. However, this information was only analysed to assess the impact that the number of asymmetries had on injury risk, and as such a comparison of exposure levels between injured and non-injured players was not provided. It would be fair to assume that the higher exposure level an athlete has in a season, the more at risk of injury they are, no matter their total FMS composite score. By including exposure data within a study, it allows injury data to be normalised, and would help to improve the accuracy of any analysis relating to injury and FMS scores. The current research makes no allowance for the differences in exposures that may exist between players above or below the cut-off score of 14. If differences do in fact exist, they will undoubtedly lead to inaccuracies in the reported links between FMS scores and injuries, if the exposure data is not included. By considering a players FMS score, whilst also considering their exposure level, studies would be able to better understand if correlations did exist between a players FMS score and their injury occurrence levels.

Exclusion Criteria

Another limitation of the current research around FMS and injury prediction is the fact that the exclusion criteria for participating in studies varies greatly, particularly in relation to previous injury incidence. No reference to the exclusion criteria was made in the studies by Kiesel et al. (2007) and Peate et al. (2007). Other studies excluded participants based on the occurrence of a recent injury. Chorba et al. (2011) and Duke et al. (2017) excluded any individual who had received an injury in the 30 days preceding testing. Appel (2012) excluded anyone with a current injury. Letafatkar et al. (2014), excluded any participant who had reported an injury within the past six weeks that was likely to affect their FMS performance. Garrison et al. (2015) excluded participants who were unable to participate in their chosen sport or averaged less than 3 hours per week sports participation. Most of these decisions were based on self-reporting of injury from the participants themselves, and as such could have led to some existing injuries not being reported. Because of such variation, there is a risk that some studies reported a strong link between low FMS scores and injury, when in fact the injury that was reported was actually a re-injury of an older injury that had previously occurred and resulted in that individual's performance on certain FMS subtests being weaker than would have been the case if the participant was fully fit.

Cut-off scores applied

Five of the studies in Table 2.6 that reported a link between FMS score and injury, used a cut-off score of 14 out of 21 to group participants in order to identify if their FMS score could be linked with injury. A score of 14 has been determined to be the minimum score required for an individual to be deemed to have sufficient movement capability to continue to participate in their chosen activity or sport (Functional Movement Systems, 2017). The study by Kiesel et al. (2007) was the first paper to use this cut-off score in their sample of 46 NFL players, and additional research continued to use this cut-off score, despite their samples being from different sports or occupations (Chorba et al., 2010; Duke et al., 2017; Garrison et al., 2015; O'Connor et al., 2011), yet still reported a link between FMS scores of 14 and below and injury risk. Therefore, this brings into question whether the cut-off scores used in research assessing FMS and injury, should take into account the physical capabilities of the sample being tested and be adjusted accordingly.

From the studies that reported links between FMS and injury, many of the samples included physically demanding occupations such as firefighters or marine officers, or highly physical sports such as American football and rugby. Participants in these occupations or sports require a high degree of physical competency to deal with the demands placed upon their bodies, thus it is likely that participants in these highly physically occupations or sports would exhibit higher FMS scores in general compared to the average individual. Duke et al. (2017) reported mean total FMS composite scores of 15.3 for their sample of 73 male rugby union players, compared to 16.9 in the study of 46 American footballers by Kiesel et al. (2007), and 16.6 in the studies of 874 marine officers by O'Connor et al. (2011) and 433 firefighters by Peate et al. (2007). These mean total FMS composite scores highlight that the majority of participants in these studies scored above the cut-off score of 14. A similar relationship was found in the studies that reported no link between FMS score and injury. Mean total FMS composite scores were 13.5 in the study of 237 junior Australian Rules football players by Chalmers et al. (2016) and 13.1 in the study of 167 high school athletes by Bardenett et al. (2015). Therefore, the majority of participants scored below the cut-off score of 14. These samples with less physically competent participants may have benefitted from lowering the cut-off score to assess if this had any impact on the results. Future research in this area should consider the physical competencies of their samples based on the demands of the occupation or sport and use cut-off scores that are relative to these physical attributes, in order to provide a more accurate indication of the link between FMS and injury.

Sample sizes

Studies that have assessed the ability of the FMS to predict injury have typically used small sample sizes for injury surveillance studies of this nature. The sample sizes for these studies ranged from a high of 874 (O'Connor et al., 2011) to a lowest sample of just 38 (Chorba et al., 2010), with seven of the studies outlined in Table 2.6 having less than 200 participants. When combined with a surveillance duration of just one season, the statistical power of these studies is low, increasing the likelihood of errors in the findings. To date, the research has not provided any indication as to the sample size required to sufficiently power the studies to increase the accuracy of the findings. This makes it difficult to draw firm conclusions from some of the research to date, as to whether the FMS scores can be used to predict injury. In order to build on the current literature, future research should consider using a large sample size, based on the statistical power calculations that provide an indication of the size of sample needed to minimise errors in the findings.

The link between FMS scores and injury is worth further exploration. However, consideration should be given to how such studies are carried out to increase the relevance of the work they complete. Understanding the effect FMS scores have on all injuries and different injury types would help to make more informed judgements regarding the role the FMS can play in injury prevention. In addition, analysing the effect subtest scores have on injury occurrence in different populations would perhaps give a stronger indication of any link that may exist between the FMS and injury. Finally, consideration should be given to the most suitable cut-off score to use, dependent on sample type, to ensure it is relative to the physical competencies of the participants and the sample size required to reach a level of statistical power that reduces the chance of error.

2.2.4 FMS SCORE AND PERFORMANCE

In addition to the research linking FMS scores and injury, further studies have concentrated on the link between performance and FMS score. Okada, Huxel, & Nesser (2011) conducted a small study on 28 recreational athletes from varied backgrounds. Participants were tested on each FMS subtest along with three performance related tests to establish if correlations existed between the two. They reported significant correlations between certain FMS subtests and each of the performance tests. The Backward Medicine Ball Throw (BOMB) was positively correlated with the HS, TSPU and RS subtests, but was negatively correlated with the SM subtest. The T-Agility test was positively correlated with the SM subtests and negatively related with the HS subtest. The Single Leg Squat was negatively correlated with the SM subtest and no correlation was found between the Core Stability test and FMS subtests. As the performance tests involved whole body co-ordination and multi-joint movement, it could be suggested that certain FMS subtests are more effective than others in predicting performance in such movements.

Chapman, Laymon, & Arnold (2013) focussed their study on the athletic population by recruiting a larger sample of 121 elite track and field athletes. Using the cut-off score of 14, participants were grouped as either Hi or Low according to their FMS total score, and their best track and field performance obtained in 2010 and 2011. Those grouped in the Hi FMS group had a significant improvement in their performance from one year to the next. This improvement was only seen for male athletes and not female. In addition, they considered the effect of asymmetries, and found those with no asymmetries improved their performance from one year to the next by a mean 0.6% compared to a 0.26% mean improvement for those who displayed at least one asymmetry. Lastly, this study considered the FMS score on a subtest (deep squat) and sprint performance and found those scoring three on this subtest had a significantly larger mean improvement in performance. These results appear to strongly suggest that total FMS scores can be used to accurately predict the ability to improve performance in track and field athletes. A limitation to this study is that it only considers the best performance from each of the 2010 and 2011 seasons for each participant. This does not take into consideration any external factors that could have affected performance such as injury, conditions, nerves etc. or consider the frequency of competition for each participant. Including a performance test protocol, introduced in both years, would have provided a more controlled environment within which to accurately test each participant and reduce the effect of external factors.

Further evidence of FMS scores and their links with performance was reported by Conlon (2013) in a small study of 36 male students. He established that countermovement jump height was significantly correlated with FMS total score using both a live and video 21-point scoring scale. However, Parchmann & McBride. (2011) reported conflicting findings in their study of 25 Division 1 golfers. They analysed if any links existed between four performance tests and the FMS composite score. They reported no correlation between FMS score and three performance tests and suggested that the FMS is not an adequate field test to determine athletic performance. The sample populations in these two studies were relatively small and also quite different when considering their physical activity levels. Waldron, Gray, Worsfold, & Twist (2014) also stated that the FMS should not be used to measure athletic ability based on the results of their small study of 12 elite level U19 rugby league players. Although this study did not compare the performance of players based on FMS subtests, it did measure season long improvements in FMS scores and various fitness components. Whilst significant improvements in all fitness components were recorded, no such improvement was seen in the FMS total scores of the players. These findings could be because of fundamental movement not improving on a consistent basis across all individuals within the sample group. Alternatively, it could be suggested that the FMS scoring system itself may not be sensitive enough to pick up small improvements in fundamental movement that occur as a result of a training programme to improve physical attributes. The developers of the FMS (Cook et al. 2006) have themselves suggested that physical performance and fundamental movement, as measured by the FMS, are two separate constraints. As such one can be developed without necessarily affecting the other. Given the small sample sizes, further work in this area would be warranted to understand the role, if any, the FMS can play in helping to improve physical performance through the development of suitable training programmes.

2.2.5 INTER-RATER AND INTRA-RATER RELIABILITY

For a test to be effective in an applied setting, it must have the ability to be performed repeatedly in an accurate and reliable manner each time it is conducted by the same tester, and by different testers. Minick et al. (2010) undertook a study to compare the reliability of two novice and two expert testers when scoring the FMS. They each tested 40 healthy college students and the results suggested the FMS had a high inter-rater reliability and could confidently be applied by trained individuals using the approved scoring system. Agreement levels ranged from moderate to excellent between the two novices and two experts, and from substantial to excellent when comparing the two novices against the two experts. It should be noted that the definition of a novice in this study, was a person who had undertaken the FMS certification course and had up to one year's experience of FMS testing. This study is therefore suggesting a level of training and experience is required before the FMS can be used reliably in the field. The authors acknowledged that further studies comparing real time inter-rater reliability should be conducted instead of using video footage to score the tests. In applied settings, the use of video is both time consuming and unpractical at times, and is therefore unlikely to be used often.

Teyhen et al. (2012) completed a study on 64 active service members (US military personnel who were involved in a military training program) to establish the inter-rater reliability of the FMS between novice testers. They reported that the total FMS composite score demonstrated moderate to excellent inter-rater reliability between two novice testers, who had received just 20 hours of FMS training prior to the study. This study suggests that even with limited training and experience, the FMS can be used reliably. Additional research provides support for these findings. Schneiders, Davidsson, Hörman, & Sullivan (2011) found inter-rater reliability for total FMS composite scores to be substantial to excellent with Intraclass correlation co-efficient (ICC) reaching 0.97 in a large study of 209 active individuals, when comparing two novice testers with limited training and experience of the FMS. More recent research has continued to identify the same trend. In a small study of 30 male ice hockey players, Parenteau et al. (2014) reported that the FMS is a reliable screen for this population. Reliability between four certified FMS testers, was particularly strong (ICC=0.96), when comparing total FMS composite score (out of 21) given by each tester. Leeder, Horsley, & Herrington (2016) also reported high inter-rater reliability (ICC=0.91) between 20 physiotherapists with no formal FMS training, when scoring five participants on six of the FMS subtests using video footage of subtest performance. When considering inter-rater reliability between different levels of testers, Gulgin & Hoogenboom (2014) reported that the ICC for mean FMS composite scores was 0.88. This indicates good to excellent consistency between one expert and three novice raters, and further suggests that the FMS can provide accurate total FMS composite scores no matter the experience level of the tester.

However, Shultz et al. (2013) in a study of 39 college athletes, found inter-rater reliability between six testers was poor. This included a comparison of two testers with less than one-year experience (ICC=0.44) and four testers with more than two years' experience (ICC=0.17), thus indicating larger differences between testers with greater experience.

When considering intra-rater reliability, Teyhen et al. (2012) found that this was not as strong as inter-rater reliability, with only a moderate to good relationship (ICC=0.74) identified at 48 hours and 72 hours, test to retest, in their study comparing two novice testers. Shultz et al. (2013) compared intra-rater reliability in a test-retest situation and between live and video scoring to ascertain if one method was more reliable than the other. They reported that the reliability between live and video scoring was consistently excellent (ICC=0.92) and was good (ICC=0.6) for test-retest reliability. Parenteau et al. (2014) also reported excellent intra-rater reliability (ICC=0.96) for video raters scoring the same video with a six-week break in between ratings. Intra-rater reliability would appear to be weaker when using a live test versus re-test protocol compared to using a live test versus video protocol. The ability to slow movement down, re-observe performance and undertake analysis away from the field, are all possible contributory factors to this improved reliability using video scoring.

Intra-rater reliability for total FMS composite scores appears to be excellent when using video scoring to assess test performance. The practicalities of this method do limit its use within applied settings, particularly given the time required to analyse video footage accurately. Such good agreement in total FMS composite scores can mask poor agreement in subtest scores, simply because these differences are averaged out across the seven tests. Table 2.7. shows perfect agreement in total FMS composite scores between three hypothetical raters, despite there being no perfect agreement on any of the seven subtests between the three raters. Although a rater only has the option of giving one of four scores on each subtest, an incorrect score on any one subtest, can result in an ineffective intervention programme being supplied to that individual. A practitioner would normally evaluate an individual's results by looking for any scores of zero or one, and particularly those with an asymmetry, on a particular subtest and provide follow up action based on this information accordingly. It would be unlikely that they would simply take the FMS composite score out of 21 and use that to prescribe an intervention programme. As such, any errors between ratings of subtest performance could have significant consequences for the individual's development moving forward.

Table 2.7. Example of hypothetical situation showing perfect agreement in total FMS composite scores between three raters despite moderate agreement between subtest scores

	Rater 1	Rater 2	Rater 3
Deep Squat	2	2	1
Hurdle Step	1	2	3
Inline Lunge	3	2	3
Active Straight Leg Raise	1	2	2
Shoulder Mobility	2	1	2
Trunk Stability	3	3	2
Rotary Stability	2	2	1
Total Score	14	14	14

Whilst inter-rater reliability compares the level of agreement between different testers when scoring the FMS, intra-rater reliability considers agreement within one tester when scoring the FMS on different occasions. In an applied setting, it is often the case that one individual will be responsible for undertaking pre-performance testing, and as such the accuracy of scoring by the same tester across different screening sessions could be critical. In addition, the inter-rater reliability of the individual subtests is particularly relevant to the programme that is prescribed to an individual following completion of the FMS. The FMS Manual (Functional Movement Systems, 2017) provides an algorithm outlining how subtest scores should be used in a particular order to develop suitable exercise programmes. It is therefore imperative that the accuracy of subtest scoring is high to ensure the correct intervention or programme is prescribed to the individual screened.

In regard to subtest scores and inter-rater reliability, the relationship is not as strong as that seen when using total FMS composite scores. Minick et al. (2010) reported noticeable differences between pairs of raters and individual subtest scores. Agreement was still no lower than moderate between subtests, but the authors suggest that the inline lunge, hurdle step and rotary stability subtests, have less clearly defined descriptors of mid-range performance making division of intermediate scores less apparent. The inline lunge and rotary stability subtests also recorded the lowest inter-rater reliability in the study by Shultz et al. (2013) with only 0.10 and 0.25 kappa values respectively, compared to the highest kappa value of 0.95 for the hurdle step subtest, with perfect agreement having a value of one. Similarly, Gulgin & Hoogenboom (2014), when comparing subtest scores between raters in their study, found the level of agreement reduced substantially for the deep squat and rotary stability subtests, to just

33% agreement using the Fisher's Exact test. This compared to 100% agreement between raters in the active straight leg raise and trunk stability subtests. Interestingly the percentage agreement for the hurdle step, inline lunge and shoulder mobility subtests was 100% when measuring one side of the body, and only 66% when measuring the other side. The authors suggested that a reason for the reduced agreement across the various subtests is a result of the scoring system available to testers, which could lead to them having difficulty in discerning where movement failure occurs, and in describing or quantifying such failures. The rotary stability subtest also demonstrated only fair reliability in the study by Parenteau et al. (2014) with a 0.26 kappa value, which was significantly lower than the 0.99 value for the shoulder mobility subtest, 0.97 value for the active straight leg raise subtest and 0.90 for the trunk stability subtest. These authors also mention the difficulty in reliably scoring certain subtests because the criteria for some are more subjective than that of other subtests. They suggest differences may exist between what two evaluators consider as an acceptable compensation in subtests such as the hurdle step, as there are no induced compensations as in the deep squat subtest, where the FMS 6ft x 2ft board is introduced if a three is not achieved. Additional research by Whiteside et al. (2014) suggested that manual assessment of the FMS was not a valid measurement unit due to the subjectivity involved in the scoring system. The authors attempted to develop an automated measurement unit, using inertial measurement devices, that could provide an objective measure of six FMS subtests using a small sample of 11 participants. Comparisons between the automated system and manual assessment by a certified FMS tester showed poor levels of agreement and Whiteside et al. (2014) suggested that the ambiguity of the scoring criteria provides a major challenge for testers and increases the subjectivity of the scoring system.

The review of literature in this area has provided inconclusive results regarding the ability of the FMS to predict injury in various populations. Whilst there does appear to be some relationship between the two, the research to date does not conclusively show that an individual's FMS total composite score can be used to predict their likelihood of receiving any type of injury. This is partly due to the limitations in the study designs employed in several studies that make it difficult to draw any conclusions from the research undertaken. Additionally, the relationship between total FMS composite score and performance is also unclear from this literature review. The small number of studies, and the low sample sizes used in most of these studies, mean further work in this area is needed before conclusions regarding the relationship between total FMS composite score and performance can be made. Finally, the research highlights that the FMS is a reliable test with good levels of inter-rater and intra-rater reliability when considering the total composite score. However, this level of reliability reduces when considering subtest scores, which are an important component of the overall testing process, and as such suggest that improvements to the FMS scoring system or the way the FMS is assessed would be worthwhile.

2.3 MOTION TRACKING TECHNOLOGY

Human motion capture consists of the recording of human body movements for immediate or delayed analysis and playback (Kolahi et al., 2007). This area of research has become increasingly popular, particularly from an athletic performance perspective, where segmenting part of the human body, tracking the movement of joints over an image sequence and recovering the underlying 3D body structure add significant value to practitioners (Aggarwal & Cai, 1998). Traditionally motion analysis of the human body has involved low level processing, such as body part segmentation, joint detection and identification, and the recovery of 3D structure from the 2D projections in an image sequence (Aggarwal & Cai, 1998). These systems have become increasingly sophisticated in recent years as technological advances have been utilised to enhance the capabilities and efficiencies of both the hardware and software elements within each system.

2.3.1 AUTOMATING THE FMS SCORING

With the reliability of scoring the FMS subtests still inconclusive, there was value in exploring the option of the development of a human motion analysis system to provide an automated assessment of the seven FMS subtests. Automating the FMS scoring process has strengths and weaknesses, which were considered before attempting to adopt an automated approach. The FMS has been designed to be a manual scoring process, which can be completed using a basic testing kit, in an efficient manner, using a simplistic scoring system (Cook et al., 2006). It is designed to be a diagnostic tool that helps to direct future interventions, testing or treatment, rather than a screen that will provide the practitioner with immediate answers regarding an individual’s technique and how one might improve areas of weakness.

One of the main challenges to automating the FMS is the scoring criteria currently employed. This involves a simple scoring scale ranging from 0 to 3 for each subtest. A score of zero is given if pain is felt and a score of three is given if performance meets the criteria provided in the FMS Manual (Functional Movement Systems, 2017). If performance includes a level of compensation, a score of two is given and a performance that involves large levels of compensation, or results in a failed attempt, results in a score of one being given. Whilst this scoring system is simple to follow and administer, it does not provide any biomechanical joint angle or alignment information that can be integrated easily into an automated system. Therefore, an automated system to score the FMS would either need to develop its own scoring criteria or attempt to develop joint angle and alignment measurements based on the existing FMS scoring criteria.

To date, two studies have attempted to automate a subset of the FMS subtests using Inertial Measurement Units (IMU's) to automatically assess the FMS. As previously discussed earlier in this chapter, Whiteside et al. (2014) reported poor levels of agreement between their automated IMU system and manual assessment of six of the seven FMS subtests in their study of 11 subjects. Although these authors suggested that this low level of agreement was caused due to the subjectivity and ambiguity of manual assessment, they also accepted that the thresholds that would have been required to be set for the IMU system would have reached unreasonable levels for some subtests in order to improve agreement levels. Additional research by Jensen, Weilbrenner, Rott, & Eskofier, (2013) also developed a semi-automated IMU system to assess the deep squat FMS subtest. This small study involving 10 subjects, showed good agreement between the semi-automated system and manual assessment, of 80%. However, certain key variables involved in the performance of the deep squat were not measured by the IMU system, such as alignment of the knee and feet in the frontal plane, which was a major limitation of this study. The authors also acknowledged that the use of the IMU system would not be possible for subtests such as the active straight leg raise and shoulder mobility, due to the intricacies and manual measurements involved in these subtests. Both studies attempted to incorporate the existing FMS scoring criteria into their automated systems, rather than creating their own scoring parameters, which as both studies alluded to, provided some major challenges in regard to reaching the level of reliability and accuracy required for each subtest. However, from an applied perspective this would be the preferred approach, as the aim of an automated system is to enhance the reliability of the current FMS scoring, rather than trying to change how each of the subtests are assessed. The FMS is a tool that forms part of a wider functional movement system and is the entry level screen that allows practitioners to examine fundamental movement patterns. The results of the FMS dictate future course of action and which test to undertake next in order to learn more about each individual’s movement competency. By changing the scoring criteria simply to suit the development of an automated system, the ability to fit the FMS into this wider functional movement system approach would be compromised.

Whiteside et al. (2014) suggested that the manual scoring system was not a valid measurement tool against which to compare an automated system, due to its subjective nature. This highlights a major weakness of the FMS and outlines why an automated approach could be of benefit, despite the challenges faced in developing such a system. Whilst FMS scoring remains a manual process, it will always be susceptible to rater error leading to inconsistencies in the scoring system. As previously highlighted this leads to poor levels of reliability when considering subtest scoring (Shultz et al., 2013; Gulgin & Hoogenboom, 2014; Parenteau et al., 2014), which form a key part of the decision making process in regard to future corrective exercise programming to enhance fundamental movement patterns. This level of subjectivity will always be apparent whilst a simplistic scoring criteria is in place and the process is conducted manually, as the entire assessment of performance is reliant on the raters interpretation. However, if this scoring criteria were to become more quantitative in nature, it would completely alter the concept of the FMS as a screen that is a simple, cost effective and time efficient tool. Adding joint angle or alignment calculations to the assessment criteria will make the screen a great deal more complex, requiring additional tools to be used and make the FMS a less practical tool in an applied setting. Therefore, an automated system that can replicate the current scoring criteria, yet reduce the level of subjectivity seen in manual scoring, would be a useful tool for practitioners who could be confident in the level of accuracy and reliability of the scoring system.

A further consideration when developing an automated approach is to ensure that the system is practical and simple to use (Mundermann et al., 2006). In applied settings, practitioners are under pressure to deliver results and their time is often a precious commodity. One of the current strengths of the FMS as it currently exists is the speed of set up and test assessment (Bardenett et al., 2015), which allows practitioners to undertake testing on a squad of players in a reasonable amount of time. There are no requirements for cameras, markers, goniometers, inclinometers etc, as the practitioner has all the tools required when using the FMS kit and following the FMS scoring criteria. Whilst an automated system would obviously involve additional items or tools to be incorporated during testing, the technology chosen needs to be as non-intrusive as possible and allow the screen to be set up and completed in the same time as manual assessment. Additionally, the automated system must be portable but with the ability to be set up in a reliable manner for every testing session. It is unlikely that a practitioner will have the luxury of being able to complete the FMS in the same location for every single assessment and it is more realistic to expect screening to take place across various locations within an organisations facility. Therefore, the automated system must be able to be moved on a regular basis, but without this having a negative effect on its level of accuracy or reliability.

A final requirement of an automated system is the ability to make the corrective exercise prescription process automated as well as the FMS scoring. Often data analysis and development of a suitable intervention is the most time consuming part of the FMS which can lead to practitioners deciding not to use the FMS in order to save themselves significant time. It is not enough for an automated system to just be able to accurately and reliably score the FMS subtests. It will also need to be able to automatically interpret the subtest scores, accurately select the necessary corrective exercises or course of action and be able to deliver this information in a format that is clear, concise and practical to the practitioner, athlete or both. This will require bespoke software to be developed, which will be based on the Corrective Exercise Algorithm outlined in the FMS Manual (Functional Movement Systems, 2017). Similarly to the scoring criteria, this algorithm has been designed with manual scoring in mind and will require the developers to interpret its contents in order to attempt to replicate it in the automated system. Whilst this presents a challenge, there would be limited benefit in developing a new corrective exercise algorithm specifically designed for the automated system, as this would effectively change the post screening process, which has been developed based on empirical evidence and research.

The previous paragraphs have outlined the challenges faced in developing an automated system to assess the FMS, the success of previous attempts to develop automated approaches and why there is justification for reducing the scoring subjectivity through such an automated tool. The following paragraphs consider how motion analysis techniques has developed in recent years and discuss the types of technology that could be used to develop an automated system to assess the FMS, focussing on the strengths and limitations of each based on the challenges and requirements highlighted in previous pages.

2.3.2 DEVELOPMENTS IN MOTION ANALYSIS

As the demands placed upon athletes has increased, the requirement for real time feedback to enhance performance has also grown significantly. As a result, human motion analysis has faced many challenges as attempts have been made to improve the accuracy and speed of 2D and 3D systems. These include:

Segmentation

Segmentation relates to the process of separating regions of the human body from the rest of an image when using motion analysis techniques. Fast and accurate motion segmentation has been a significant and often difficult problem (Wang, Hu, & Tan, 2003). The captured images in dynamic environments are often affected by factors such as the weather, lighting, shadows and camera motion. This has meant motion segmentation has had to develop more reliable models adaptive to dynamic changes in complex environments.

Occlusion handling/multiple cameras

Historically, human motion analysis capture systems have had issues with self-occlusion of human body and mutual occlusions between objects (Wang et al., 2003). Typically, during occlusions, only portions of each person are visible and often at low resolution. Primarily this has been an issue for single camera systems. Thus, multiple camera solutions have been developed to alleviate these issues. The availability of information from multiple cameras can help to expand the surveillance area and provide multiple viewpoints, compared to a single camera viewpoint which can easily generate ambiguity due to occlusion or depth (Wang et al., 2003).

Performance evaluation

Robustness, accuracy and speed are three major demands of practical human motion analysis systems (Moeslund & Granum, 2001). Systems need to be robust and insensitive to noise, lighting, weather etc. as well as having the robustness to cope with large amounts of data and a number of different users (Wang et al., 2003). The accuracy and speed of data collection/processing is also a major consideration for human motion analysis systems and an area that has seen major developments in recent years. This allows systems to provide real time feedback with a high level of accuracy, enhancing the effectiveness of the outputs produced by each system and their usefulness in applied settings.

As systems have developed to meet these challenges, there are several options available regarding suitable technology to adapt and utilise for the purpose of assessing the seven FMS subtests. The following paragraphs will review three different types of systems which offered a potential solution to reduce the subjectivity of the FMS scoring system.

2.3.3 INERTIAL MEASUREMENT UNITS

Inertial measurement units (IMU's) are self-contained, non-radiating[1], non-jammable[2], dead reckoning[3] devices that provide dynamic motion information through direct measurements (Altun, Barshan, & Tunçel, 2010). They often contain either gyroscopes that provide angular rate information around an axis of sensitivity, or accelerometers that provide linear or angular velocity rate information (Altun et al., 2010), or feature both pieces of technology. In the past 10 years the size, weight and cost of commercially available sensors has decreased considerably, which has opened up new possibilities for the use of IMU's in human activity monitoring, recognition and classification through body worn sensors (Altun et al., 2010).

Research on IMU's has focussed on their ability to accurately track various movements such as stride parameters, single limb movement and whole body movement. Tan, Wilson & Lowe (2008) assessed the accuracy of an IMU combined with a low cost and high cost GPS unit to determine stride parameters during running. Although a limited sample size of just one participant was used for the study, the authors reported that integrating the low cost GPS unit with the IMU improved measurement of running speed,

---

[1] The signal from the device cannot be disrupted from frequencies emitted from other devices

[2] Device does not emit electromagnetic radiation

[3] The data collected from an IMU allows a PC to calculate the devices positioning

and in particular the estimation of stride frequency and mechanical energy fluctuations. Further work in the area of gait analysis compared the accuracy of an inertial sensor system against an instrumented treadmill to asses joint kinematics and spatiotemporal parameters using 20 participants (Nuesch, Roos, Pagenstert, & Mundermann, 2017). The IMU's were placed on the sacrum and bilaterally on the lateral thigh, lateral shank and lateral foot. The authors found there was excellent coefficients of multiple correlation for all joints (between 0.97 and 0.99) and that the root mean square error between the waveforms measured by the two systems, after offset correction, was smaller than 5° for all joints. They concluded that the joint angles measured in the sagittal plane were highly correlated, but only after offset correction, and that the test-retest measurements for the inertial sensor system were very good to excellent.

Junker, Amft, Lukowicz, & Tröster (2008) presented a novel two stage gesture spotting method based on body worn IMU's. Their method was tested on four participants by attaching five inertial sensors on the wrists, upper arms and upper torso, and assessing if various arm gestures had occurred during recording of continuous, non-task specific, data. They demonstrated that their strategy could feasibly be used for the identification of motion events in a continuous signal stream and worked well for arm based motions that are difficult to recognise due to the inherent complexity of arm motions.

Additional work focussing on arm movement was conducted by Zhou, Stone, Hu, & Harris (2008). Using two inertial sensors attached to the upper and lower arm of each of the four participants, they tracked the movement of the arm using the IMU's whilst each participant completed a "designed paths" experiment that included a circular and then a rectangular movement as instructed by the authors. Following the first task, one sensor was placed as close to the wrist centre as possible and the second on the lateral aspect of the upper arm. Each participant completed three additional tasks; target reaching, shoulder shrugging and forearm rotation. To assess accuracy of the IMU's, a CODA motion capture system was used to track arm movement during each of the tasks. The authors reported that the IMU's had a high level of accuracy in tracking arm movements compared to a motion capture system, with root mean square position errors that were less than 1 cm and root mean square angle errors of 2.5 to 4.8º. However, the authors did report that the IMU's failed to accurately detect smaller movements e.g. less than 0.5 cm or 2º, which they attributed to the relative movement of the sensors on the arm.

The research outlined in previous paragraphs shows promise regarding the use of IMU’s to assess various aspects of human movement. However, the research in this area has outlined limitations to IMU’s that provide challenges to incorporating them as part of an automatic solution to assess the FMS subtests. One such limiting factor is the amount of drift that an IMU can suffer from, which can negatively affect their application in poorly controlled environments (Zhou et al., 2008). Drift is essentially an ever-increasing difference between where the sensor thinks it is in terms of location, and where its actual location is. Traditionally such errors build up over time to the point that the IMU become unusable. From a practical point of view, with any automated system designed to assess the FMS needing to be robust, durable and reliable, any potential long-term errors that could cause the system to malfunction would be unacceptable. A further difficulty with the use of IMU’s is the limited guidance that exists on finding a suitable configuration, number and type of sensor to use (Altun et al., 2010). Across the literature such variables differ widely among studies, along with the signal processing and motion detection techniques employed. Added to the fact that many of the studies in this area have used very small sample sizes, as outlined in the previous paragraphs, it makes comparisons between studies very difficult and often meaningless.

Junker et al. (2008) identified a further limitation in their study, which was the difficulty an IMU has in spotting sporadically occurring activities in a continuous data stream. This is particularly relevant for applications that are attempting to monitor specific tasks or movements, as these often occur sporadically in between a large variety of other activities. Considering the multi-trial and multi-joint requirements of each of the seven FMS subtests, the ability of an IMU based system to manage and interpret data to provide an automatic score, would be a concern with this identified limitation. Finally, an obvious limitation of using IMU’s is the need to accurately place the sensor on the participant’s body repeatedly time after time. Due to the fact such placement is conducted by a human tester, the calculations used to determine sensor location are prone to error. Zhou et al. (2008) suggested that IMU errors are often caused by the relative movement of sensors against the underlying bony anatomy or incorrect sensor placement due to the incorrect identification of joint centres. This could lead to inaccurate data being captured as the exact gesture or movement being tracked could be different each time merely due to sensor placement. The practicalities of securely placing a sensor on different body parts within a restricted time period will offer many challenges if an IMU system was to be implemented for automatic assessment of the FMS.

2.3.4 MARKER BASED SYSTEMS (MBS)

These systems typically employ the use of small markers attached to the body of the subject and a set of two or more cameras focussed on the subject to capture its motions (Kolahi et al., 2007). The vast majority of marker based systems model the limb segment as a rigid body through the application of markers to specific body parts or joints, then apply various estimation algorithms to obtain an optimal estimate of the rigid body motion (Mündermann, Corazza, & Andriacchi, 2006). These systems are often used for personalised training systems for various sports, medical diagnostics, and choreography related activities (Wang et al., 2003), as well as for research purposes in academic institutions.

A MBS normally requires precise set up prior to each use, with a thorough calibration required of the cameras to ensure accuracy of measurement and consistency compared to previous data capture. Performance of a MBS strongly depends on their setup and is highly sensitive against alterations (Morlock, Windolf, & Gotzen, 2008). As a result, these systems are most commonly found in controlled laboratory based environments to help protect the integrity of the system and increase the likelihood of achieving reliable and valid data collection. The popularity of MBS has increased as the technology has improved, and more commercial systems have become available. Companies such as Vicon, Qualisys and Optotrack have led the way in developing these systems, and such development has been supported by research in the field, assessing the reliability and validity of MBS. As a result these visual marker based tracking systems are quite often used as a “gold standard” in human motion analysis due to their accurate position information (errors are around 1mm) (Zhou & Hu, 2008).

Although MBS are widely used across several industries and are accepted as a gold standard method by which to assess human movement, they do provide some challenges from a practical perspective. One such challenge is that the cost of an MBS can be extremely prohibitive. These systems typically range from a few thousand pounds to much larger sums for systems that integrate the motion capture system with instrumented treadmills and other high tech scientific measurement devices. A further challenge to using a MBS in real world applications is the complexity, bulk and space requirements involved (Dutta, 2012). A MBS typically needs to be set up in a very controlled environment (Fernández-Baena, Susín, & Lligadas, 2012), with the cameras located in a consistent manner and requiring a calibration process at the start of each testing session. In addition, a certain amount of space and equipment is required in order to set the cameras up correctly, which provide significant challenges from a practicality point of view when using in applied settings. For the purposes of automatic assessment of the FMS, the cost and practicality of a MBS makes it an extremely difficult option despite some of its advantages.

From an accuracy and reliability perspective, several authors have highlighted a major issue with MBS. This is specifically related to how these systems calculate the joint centres and the impact of soft tissue artefact. Taylor et al. (2005) suggest that MBS are subject to soft tissue artefact, where the attached markers move relative to the underlying skeletal structures, resulting in errors in the determination of the skeletal motion. These authors highlight that subjects with high tissue coverage will be more at risk of incorrect analysis because of this limitation. Additional research has also reported concerns in regard to soft tissue artefacts (Herda, Fua, Plankers, Boulic, & Thalmann, 2001; Mündermann et al., 2006; Zhou & Hu, 2008) and the impact this can have regarding the accurate assessment of motion. Attempts have been made to reduce the impact of soft tissue artefact through adopting approaches such as the Point Cluster technique (Alexander & Andriacchi, 2001) and Optimal Common Shape Technique (Taylor et al., 2005).

The point cluster technique (PCT) employs an overabundance of markers (a cluster) placed on each segment to minimize the effects of skin motion artefact (Alexander & Andriacchi, 2001). These authors conducted a study to develop and test an extension of the PCT by correcting for error induced by segment deformation associated with skin marker movement relative to the underlying bone. They performed a simulation study that involved running 50 trials based around an eight-marker cluster set. Upon successful completion of the simulation an in vivo study was performed on one participant. The subject was fitted with an Ilizarov external fixation device, for the purposes of tibial lengthening. The point cluster marker set was affixed to the subject’s shank (six markers), along with a set of four markers rigidly attached to the Ilizarov device, which was connected to the tibia with bone pins. These four markers define a true bone embedded coordinate system. The device was rigidly attached to the tibia with nine bone pins, located in three sets of three pins. The subject performed a 10 cm step up task. Results indicated that this technique reduced the average location error from 0.03 to 0.01

cm and the average orientation error from 0.37 to 0.08 per time step, suggesting that this process can significantly reduce the amount of error associated with skin marker motion.

Whilst the PCT has shown initial promise, it remains to be independently verified relative to bone kinematics, particularly for the femur where the soft tissue envelope is much greater (Taylor et al., 2005). In their study, Taylor et al. (2005) proposed an alternative method for reducing the influence of individual marker movements by forming an optimal configuration of the markers throughout the measured timeframes. They termed this alternative method the Optimal Common Shape Technique (OCST), which would effectively reduce any individual marker fluctuations, without the requirement for subjective reference frames, but not address the synchronous shifting of an entire marker set relative to the underlying bone. Using three sheep, fitted with three schanz screws to the femur, tibia and metatarsus of the right hind, they assessed various gait parameters. Each screw had a light aluminium frame that held four markers attached to it, whilst four to six reflective markers were attached to the skin of each body segment. Each sheep was walked over a gangway between two and ten times, whilst a motion capture system recorded the marker positions. Results indicated that the OCST produced a small improvement on skin marker placement compared to PCT, but such improvement was minimal. The authors suggested that the limb segments with larger amounts of soft tissue coverage consistently produced larger movement errors, and implied that any improvement of techniques should improve upon factors such as the recognition of subject specific soft tissue coverage or the active identification of underlying muscle firing. They also reported that the errors associated with individual marker fluctuations through skin elasticity and non-rigid marker movement only played a partial role in the cause of skin marker errors. The errors were thus also associated with the unison movement of the marker set, most likely caused either by the swelling of tensioning muscles or the impact of the hoof on the ground.

Additional methods to functionally determine spherical joint centres, using the symmetrical centre of rotation estimation (SCoRE) (Ehrig, Taylor, Duda, & Heller, 2006) and joint axes, using the symmetrical axis of rotation approach (SARA) (Ehrig, Taylor, Duda, & Heller, 2007) have been proposed. Ehrig et al. (2006) presented the novel SCoRE technique which is a two-sided approach capable of considering a moving centre of rotation, as opposed to a one-sided approach that requires a static centre of rotation. The authors demonstrated that most commonly employed techniques can determine the centre of rotation to within 0.3 cm, if the range of motion of the joint was 45º or more. Under more limited range of motions, however, the differences in accuracy between approaches became much more apparent, but that the SCoRE technique maintained small errors even at 20º. In additional work, Ehrig et al. (2007), discussed the need to understand the axis of rotation rather than just the joint centres to accurately describe joint kinematics. They proposed the SARA technique that can consider the rotational movement of two body segments independently of each other. This technique avoids the need to transform the co-ordinates of one body segment into the co-ordinate system of another, which can be affected by artefacts, hence amplifying any inaccuracies in calculating the axis of rotation. They reported that the SARA technique can produce the most accurate axis of rotation estimates across a range of tests incorporating a virtual hinge joint with marker placements computed automatically.

Despite these advancements in reducing the impact of skin artefacts and joint centre calculation, MBS are still prone to human error in regard to marker placement, particularly when different operators are involved in testing procedures within the same laboratory. MBS normally involve placement of a minimum of 18 markers for lower limb assessment (and can involve over 70 markers for full body assessment). The various techniques and guidelines outlined in previous paragraphs are aimed at improving the consistency of marker placement and, as a result, the accuracy of data collection, but often a limitation of these approaches are the time consuming placement of additional markers (Mündermann et al., 2006). The nature of MBS means it is impossible to avoid human involvement in marker placement. Coupled with the fact that marker placement takes time to ensure each marker is placed accurately, and the chances of incorrect placement increases as operators can often be placed under pressure to test several participants in a short space of time.

Based on the limitations of MBS outlined in previous paragraphs, it appeared that this type of system would not be a suitable solution to use for developing new technology for the automatic assessment of the FMS. The practical issues of having to use a laboratory setting, with rigid calibration processes and expensive equipment, along with the prospect of accuracy issues due to incorrect marker placement and skin artefacts, made a MBS an unsuitable option for the purposes of this study.

2.3.5 MARKERLESS SYSTEMS

Measuring human movement in its natural environment, using a non-invasive method, is the most effective approach to accurately assess 3D joint kinematics (Mündermann et al., 2006). In their review of markerless motion capture Mündermann et al. (2006) stated that the development of techniques for human body kinematics estimation that does not require markers placed on the body would greatly expand the applicability of human motion capture. Not only does such a system provide a non-invasive approach, but it also helps to significantly reduce the set-up time of human motion capture and open the possibility of analysing multiple subjects at the same time. In addition, removing the requirement to place markers on the subject to calculate joint centres, removes the issue of skin artefacts and marker movement, which as highlighted in previous paragraphs, is an issue that can reduce the accuracy of marker based systems (Taylor et al., 2005).

One such markerless system is the Microsoft Kinect sensor which was developed to support the Microsoft Xbox games console to provide players with the opportunity to play games without the need for handheld controllers due to the Kinect's integrated 3D camera capabilities (Muller, Ilg, Giese, & Ludolph, 2017). The sensor integrates a Red Green Blue (RGB) camera with an infrared depth sensor to allow tracking of human movement in three dimensions without the need for markers to be attached to the individual analysed. The sensor can capture full body motion, as far away as 4.5m, as well as having facial and voice recognition capabilities. The infrared projector is combined with a monochrome Complementary Metal-Oxide Semiconductor (CMOS) sensor, allowing for video data to be captured in 3D, under any light conditions.

Two versions of the Kinect sensor have been developed by Microsoft. The Kinect v1 relied on the recognition of reflected infrared patterns to acquire the depth information. In contrast to this, the Kinect v2 used time of flight measurements, was less sensitive to interference with other sensors and provided a higher resolution (Muller et al. 2017). The term "time of flight" describes the method to determine the distance to an object by measuring the time a laser pulse needs to travel from the sensor to the object and back. The Kinect v2 provided five video data streams. Besides the colour (1920x1080@30Hz) and infrared (512x424@30Hz) data streams, it provided depth images (512x424@30Hz), body index images (512x424@30Hz) and the skeleton information for every tracked individual (25 joints@30Hz) (Muller et al. 2017). The sensors tracking volume was defined by the field of view (70º horizontally, 60º vertically) and the range of depth sensing (0.5±4.5 metres). These data streams could be accessed using Microsoft's software development kit (v2.0). Colour images were provided with four bytes per pixel and depth images with two bytes per pixel resolution. Due to its low cost and portability it potentially offered an alternative to manual assessment of the FMS through the development of new software linked with the sensor.

Several studies have been conducted to assess the Kinect’s accuracy compared to an industry gold standard MBS. To date results have been mixed and largely depend on the type of movement and body part being analysed. Bonnechère et al. (2014) assessed 48 healthy adults when completing four primary movements in one plane of motion using the Kinect sensor to assess inter and intra-rater reliability compared to the MBS. Joint kinematics were recorded whilst participants performed each task using both upper and lower limbs over two different sessions. Test-retest reliability was moderate too good for both systems based on Intraclass correlation co-efficient. However, when comparing reliability between systems, results varied, with shoulder abduction showing excellent agreement, and hip abduction and knee flexion reporting no to poor agreement.

These results appear to suggest that the Kinect is unreliable compared to the MBS when assessing joint kinematics of lower limb joints. However, as previously highlighted, some literature suggests that error levels when calculating the joint centre locations detected by the MBS can be high (Herda, Fua, Plankers, Boulic, & Thalmann, 2001; Mündermann et al., 2006; Zhou & Hu, 2008), so it is difficult to determine which method was more accurate when detecting the joint centres. The results also suggest that the Kinect can reproduce the same results in a test-retest situation as reliably as the MBS, and given its cost effectiveness and portability, further analysis of its other capabilities seemed to be warranted.

An additional study by Mentiplay et al. (2015) analysed the Kinect’s ability to assess gait parameters, including ground contact time and joint kinematics, in 30 injury free individuals. As per previous research, a comparison was made between the data produced by the Kinect and an MBS to determine the concurrent validity and inter-day reliability of the Kinect when assessing spatiotemporal and kinematic gait parameters. Two testing sessions were conducted, seven days apart, with participants performing gait trials at two different speeds; comfortable and fast paced. The results suggest that the Kinect sensor has potential to accurately assess spatiotemporal parameters but has low reliability when used to measure lower body kinematics. The authors acknowledge that the sensor placement (2.5 m away from the participant in the frontal plane) may have contributed to this low reliability finding. Further research has suggested the Kinect has the ability to be used as a home based tool by Parkinson's Disease and stroke patients to help assess normalised stride length (Cao et al., 2017), is accurate enough to measure clinical parameters of recovery stepping behaviour such as step length and time (Shani, Shapiro, Oded, Dima, & Melzer, 2017) and has the potential to be used in clinical screening programs for a wide range of patient populations (Clark et al., 2012).

The option of linking more than one Kinect sensor to improve reliability and accuracy has also been considered in recent studies. Ma et al. (2017) utilised four Kinect sensors in an ICU ward to monitor patient mobility levels. Over three hundred and sixty hours of data were collected and assessed against subjective expert analysis and a weighted kappa measure of 0.86 and weighted percent agreement of 96% were reported. The authors incorporated a sophisticated algorithm to allow the Kinect to break the recordings down into a five-step process which moved from person localisation through to mobility classification. This provided detailed feedback regarding the mobility of each patient and highlighted the capabilities of the Kinect device when incorporated with bespoke software, which can significantly increase its validity in an applied setting. Additional work, looking at multiple Kinect devices, also reported promising results when assessing spatial-temporal gait parameters when compared to a MBS (Muller et al. 2017). Using a complex algorithm to allow six Kinect's to be set up as an avenue in an overlapping method, 10 healthy subjects completed 10 x seven metre walks. The results indicated a high level of agreement between the multiple Kinect system and the MBS, especially for step length and time, and reported higher agreement for several gait parameters compared to previous research using only one Kinect system.

The research highlighted in the previous paragraphs suggested that the Kinect sensor had some promise as a tool to assess various body movements, but as yet does not seem to provide the same level of reliability as an MBS. To date, no standardised protocol has been established regarding the most effective set up of the Kinect to maximise reliability in applied settings. This includes agreement on the most appropriate height for the sensor to be set at, the most reliable distance for the sensor to be placed away from the participant, the most effective view angle and the most effective number of Kinect devices to be used. Whilst it seemed that multiple Kinects offered greater accuracy than one standalone system, this increased the complexity and cost of the set up. Additionally, the view angle seemed to have a bearing on accuracy with Muller et al. (2017) reporting that the joint positions identified by the Kinect were biased towards the surface area that is visible to the sensor.

Further research to standardise protocols relating to the Kinect set up would seem to be worthwhile given the low cost and portable nature of the Kinect sensor. In addition the markerless nature of the device allowed for increased speed of set up and removed the need for accurate placement of markers over anatomical landmarks (Clark et al., 2012). Mündermann, Corazza, & Andriacchi, (2006) highlighted that marker based systems are by nature, laboratory based, which can lead to measurement induced artefacts affecting movement such as walking gait. They also suggested the ideal measurement system should measure subjects in their natural environment and be capable of measuring motion over a sufficiently large field of view.

Although the Kinect sensor was not without its own weaknesses, particularly regarding anatomical landmark identification, it appeared to be a useful tool to help increase the speed of human movement assessment, significantly reducing the cost of measurement and allowing subjects to be monitored in their natural environment. Such positives made the device an ideal solution for the assessment of the FMS in an attempt to reduce the subjectivity of the current scoring system.

2.4 CONCLUSION

This literature review has discussed the main injury types and location of injuries recorded for soccer players. The main types of injury (muscle strains, ligament sprains and contusions) remain consistent no matter the age, gender or ability level of the players, and this trend is replicated when considering injury location, with the thigh, ankle and knee proving the most common. Due to the differences in injury definition and injury tracking techniques used across research studies, it is difficult to ascertain the most common length of injury severity. Minor and moderate injuries appear to be the most common when using the time loss definition of injury recommended by the UEFA Injury model (Hagglund et al., 2005). The review discussed several risk factors associated with injury, and found that previous injury, fatigue and exposure levels all contributed to an increased injury risk for players. Due to the multi factorial nature of injuries, each player has their own individual risk factors that are linked to body type, fitness levels, biomechanics, playing position and many other factors, which make it very difficult to develop a one size fits all approach to injury prevention and management.

This review also highlighted guidelines relating to the most effective methods to employ when undertaking an injury surveillance study. The UEFA Injury model (Hagglund et al., 2005) and the FIFA Injury Consensus Group (Fuller et al., 2006) provide clear guidelines relating to key areas such as injury definition, injury tracking, and study length, that should be considered prior to beginning an injury surveillance study. By incorporating the recommendations in these models, more critical analysis of the research in this area can be conducted, allowing a more comprehensive understanding of injuries in soccer, which can hopefully lead to practitioners reducing the high level of injuries currently occurring within the game.

The literature review provided inconclusive results in relation to the most effective use of the FMS in a sporting context. The links between the FMS, injury and performance are still questionable and thus would appear to warrant much greater research, with a clear and concise injury surveillance protocol established to guide researchers in future studies. This would allow for clearer comparisons between studies and help to assess if real links between the FMS and injury, and the FMS and performance, do exist. Understanding the impact of subtest results and asymmetries on injury and performance would also be worthy of analysis, as the use of the total FMS composite score to assess injury risk and performance may not allow for certain relationships to be identified. Lastly, the role that intervention programmes, prescribed based on FMS results, play in helping to reduce injury risk and improve performance is an area that is lacking in research at present. The role of the FMS, as set out by its developers (Cook et al. 2006) is to assess a person’s fundamental movement patterns and to identify if any asymmetries, weaknesses or compensations existed. Once these have been established the role of the practitioner is then to ascertain the next best course of action. This could include referral to medical professionals if pain is identified, prescription of a suitable corrective exercise programme to help improve the key areas of weakness, dysfunction or asymmetry, or further activity/sport specific screening to identify an individual’s suitability to participate in their chosen activity. To date very limited studies have been conducted to assess if the FMS can be used effectively in this way to improve performance or reduce injury risk, through the work completed in the intervention programme, rather than simply using the total FMS composite score.

Studies have also proven the inter-rater and intra-rater reliability of the FMS when using only composite scores, but when considering subtest scores, this reliability significantly reduces for some subtests. Although some studies have shown good agreement between raters no matter their level of experience, question marks remain regarding the level of training and experience needed to ensure subtests are carried out accurately. This is partly due to the scoring system used for the FMS, which is subjective and leaves room for error, particularly when providing mid-range scores for certain subtests. Improving the reliability of the FMS and reducing the subjectivity of the scoring system would appear to be useful objectives to further enhance the use of the FMS in applied settings.

Lastly this review has discussed a range of technologies that are currently available and could be used to develop an automated method by which to assess the FMS subtests to help improve accuracy and reliability. These methods include inertial measurement units, marker based motion capture systems and markerless depth cameras. The practicalities of using each method was discussed, along with the validity and reliability of each based on current research, to assess which technology would be the most practical to use in the context of FMS assessment. From the research undertaken, the markerless based depth camera appeared to be the most practical solution, as it was inexpensive to purchase, very portable, and research had begun to show that the reliability and validity of such technology was improving, particularly as knowledge regarding camera placement, height and angle improved, along with options to use multiple cameras to increase accuracy.

2.5 RESEARCH AIMS

This programme of research sought to build on previous research by addressing important limitations relating to injury type and injury definition, as well as analysing the impact of the various FMS subtests and asymmetries on injury occurrence and performance. In addition, the collection of exposure data, outlining the training and match minutes each player is exposed to throughout the season was completed in this research, so a clearer understanding of injury levels per player per exposure, and how these link to FMS scores, could be identified. A revised scoring system was also incorporated into the programme of research, in an attempt to remove some of the subjectivity seen in certain subtests, particularly when scoring mid-range performance. Finally, automated software linked with novel depth camera technology was developed to allow the FMS to be assessed automatically to reduce the subjectivity currently seen within the scoring system, potentially leading to improved accuracy and reliability in subtest scoring and corrective exercise prescription.

FMS and injury

The link between FMS subtests and their ability to predict injury in various populations is still inconclusive. This, in part, has been caused by inconsistencies in the methodologies used across studies relating to injury tracking procedures, injury definitions and sample sizes. There has also been little consideration given to the effects of subtest scores and asymmetries recorded during the FMS, and their potential links with injury. Finally, most studies within this area have not recorded exposure levels within sample populations, resulting in the inability to normalise injury rates. This study attempted to understand if the FMS can be used to accurately predict injury rates in soccer players through the adoption of a clearly defined injury surveillance process in line with recommended protocols (Hagglund et al., 2005). Additionally, the study analysed the impact of asymmetries and exposure levels on injury occurrence, to establish if the FMS could be used to predict injury using methods other than the total FMS composite score.

FMS and performance

To date, research linking FMS scores with performance in physical tests or sports has delivered mixed results. This could be caused by the performance measures used which have differed across studies, making it difficult to establish if a link does exist. This study assessed the link between FMS composite scores and performance tests that replicate the physical demands of soccer and the movements undertaken by players during matches and training.

Development of software to assess the FMS

The research currently suggests that there is good inter-rater reliability between raters with different levels of experience, when considering FMS composite scores. It has also been reported that intra-rater reliability is strong between live versus video scoring of the FMS, once again when considering FMS composite scores. However, such strong reliability was not apparent when looking at the FMS subtests, which could be caused by the subjectivity of the scoring system, as suggested by different authors. For the FMS to be considered a truly valid test that can be used in applied settings in a reliable and accurate manner by the same or different raters, further work needs to be completed to reduce the level of subjectivity. This programme of research attempted to develop innovative software, linked with depth camera technology, that could be used to assess the FMS subtests and provide an automated score. In addition, this research attempted to develop an automated process which will allow FMS subtest scores to be used to automatically prescribe individualised and accurate corrective exercise programmes that can be delivered in a more time efficient manner.

Validation of software to assess the FMS

As the innovative software was developed, a series of validation protocols were undertaken as part of this programme of research, to assess the validity and reliability of the software compared to manual scoring of the FMS. These protocols were used to inform the ongoing development of the software to improve functionality, with the aim of increasing the scoring accuracy throughout the development process. Upon completion of the software development, a final validation study was undertaken to assess the software’s validity and test-retest reliability when scoring all seven FMS subtests and when prescribing corrective exercise programmes, compared to manual scoring provided by certified FMS testers.

CHAPTER THREE

THE EFFECTIVENESS OF THE FUNCTIONAL MOVEMENT SCREEN IN PREDICTING INJURY RATES IN SOCCER PLAYERS

3.0 INTRODUCTION

As highlighted in chapter two, injuries continue to cause issues for practitioners working with sports teams and athletes. Injury type, location and severity have remained surprisingly constant over the past 15 years (Ekstrand et al., 2009), even though attempts have been made to improve injury prevention methods (Heiderscheit et al., 2010, Towlson et al., 2013; Halson, 2014). Within soccer, a long-term solution to injury reduction is constantly being strived for, but the multi risk factors for injury make it an extremely difficult area to improve. Each individual player has their own unique internal risks to injury that could include age, fitness levels, movement patterns, previous injury, strength and many more. In addition to these internal factors, the fact that soccer is by nature a contact sport, means the impact of external factors cannot be discounted and are very difficult to control for. This makes injury prevention a complex issue for every soccer player, no matter their ability level.

Movement screening is now common amongst professional and semi-professional soccer teams. This involves a player being screened through a series of tests aimed at assessing their movement patterns, mobility, stability and general motor control (Gulgin & Hoogenboom, 2014). Normally these would form part of a pre-season testing protocol and return to play procedure, to provide benchmark data for each player regarding their current level of performance on each test. These results are used to prescribe prehabilitation/injury prevention programmes for each player in an attempt to reduce their risk of injury and improve their performance on the pitch. The FMS (Cook et al., 2006) has gained popularity in recent years as it provides a practitioner with seven easy to administer tests that require limited equipment and provide an assessment of multi-joint movement. The FMS also helps to identify asymmetries and compensations that may exist within a player's movement, which could impact their risk of injury. To date, research has identified a link between total FMS composite scores and injury risk (Kiesel et al., 2007; Chorba et al., 2010) and links between total FMS composite scores and performance (Conlon, 2013; Chapman et al., 2014), but there are limitations to the studies that have been completed in this area, as discussed in chapter two. In regard to the FMS, there are a limited number of research studies that have been conducted on a sample population of male soccer players, and also assessed the link between additional performance tests, and FMS subtests, on performance and injury. This study attempted to assess if any relationship existed between the FMS, performance tests, and injury, using a prospective study design, including injury surveillance in line with the UEFA Injury model (Hagglund et al., 2005).

3.1 RESEARCH AIMS

This study attempted to answer the following research questions:

- Could total FMS scores accurately predict a player's likelihood of receiving an injury?
- Were there significant differences in pre-season performance test scores between players sustaining injuries and those not sustaining injuries?
- Could performance test scores accurately predict a player's likelihood of receiving an injury?
- Are total FMS scores a reliable indicator of performance on both physiological and movement based tests?

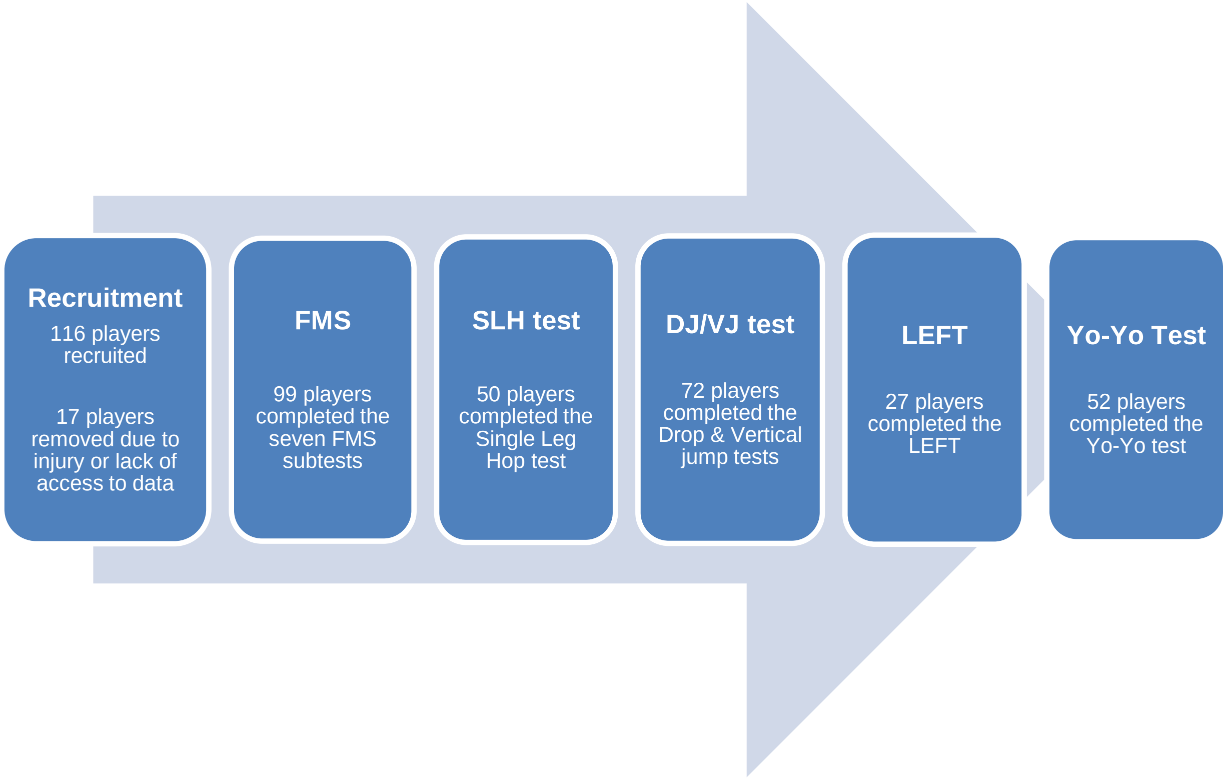
3.2 METHODS

3.2.1 POPULATION OF INTEREST

Senior League of Ireland male soccer players that were part of the first team squad for the 2014 season, aged between 18 and 40, who were fit to participate in normal soccer activities at the time of testing, took part in the study. Players were excluded from participating in the study if the club physiotherapist had declared them unfit to participate in the testing protocol undertaken by the researcher. The League of Ireland is a semi-professional league that comprises of two divisions. Player's typically train three to four times per week, as well as holding down alternative employment. Player's will often sign a one-year contract with their League of Ireland club.

3.2.2 SAMPLING METHOD

The researcher attempted to contact all twenty-two League of Ireland clubs to invite them to participate in the study. The contact details for six of the twenty-two clubs were not available and as such were not contacted. Of the remaining sixteen clubs that were contacted, six agreed to take part in the study, which resulted in 116 players being recruited. Due to restrictions put in place by the clubs involved, it was not possible to test each player on all the FMS subtests and performance tests conducted in this study (see Figure 3.1). The following paragraphs outline the tests undertaken.



FMS = Functional Movement Screen; SLH = Single Leg Hop; DJ/VJ = Drop Jump/Vertical Jump; LEFT = Lower Extremity Functional Test

Figure 3.1. Number of participants who completed each FMS subtest and Performance test

3.2.3 PILOT STUDY

A pilot study was conducted to determine the researcher’s reliability in scoring each FMS subtest and to help select the relevant performance tests for the main study. This involved screening 11 participants in the seven FMS subtests and eight performance tests during a one-off session. Each player undertook the seven FMS subtests first, and each performance was recorded using a Sony handheld camera set up 4.00 m from the participant at a height of 1.00 m. Each player was recorded completing each test twice in the frontal plane and twice in the sagittal plane. The video footage from each participant’s performance on the FMS was then sent to a certified FMS tester to score in real time. These scores were compared to the researcher’s live scoring to assess the level of agreement between the scores provided for the individual subtests and the total composite scores. The level of agreement for the total composite scores was 82% and for the individual subtests ranged from 73% to 91% agreement. These results were comparable to previous research assessing inter-rater reliability (Minick et al., 2010; Teyhen et al., 2012) and confirmed the researcher's ability to reliably score the FMS.

Each player also performed eight performance tests. These were the Yo-Yo Recovery Level 2 test, Star Balance Excursion test, Modified Thomas test, Ankle Dorsiflexion test, Lower Extremity Functional test, Single Leg Hop test, Drop Jump test and Vertical Jump test. These were chosen to provide a wide range of measures that were relevant to soccer performance including; aerobic endurance, agility, bilateral/ unilateral power and balance. Additionally, the Modified Thomas test, and Ankle Dorsiflexion tests provided a measure of the participant's quadricep flexibility and ankle range of motion. It was felt that these would be important measures when considering an individual's injury risk as these are areas of the body where injuries are common when playing soccer (Ekstrand, 2008). The Drop Jump test was selected as it provided an indication of an individual's bilateral landing mechanics. Being able to control knee valgus during such tasks has been highlighted as a possible method to reduce Anterior Cruciate Ligament (ACL) injury and patellofemoral pain (Herrington, Munro, & Comfort, 2015).

After pilot testing was completed it was decided to remove three of the performance tests from the final testing protocol. These were the Modified Thomas test, Star Balance Excursion test and the Ankle Dorsiflexion test. This decision was taken based on the time required to conduct each test and on the researcher's ability to gather accurate data from each test based on them being assessed in a consistent and reliable manner. The total testing time was over 30 minutes for each player when attempting to use eight performance tests and it was felt that this could end up negatively affecting participant numbers due to the time restrictions placed upon the researcher during the final testing protocol. In addition, the ability of the researcher to repeatedly set up and assess each test in a consistent and reliable manner was considered and the researcher was not confident in his ability to do this for the three tests that were removed. It was therefore felt that the data collected from these three tests may not be accurate and as a result unduly affect the final results of the study. This resulted in the seven FMS subtests and only five performance tests being utilised for the final testing protocol.

3.2.4 PRE-TEST PROTOCOL

The researcher visited the training facilities of each team involved in the study during the pre-season period of January and February 2014. Testing was completed during one or two training sessions, depending on the access provided by each club. Testing was completed in an indoor facility provided by each club to ensure consistency of testing from session to session. Prior to undertaking any testing, each player completed an Informed Consent form (see Appendix A) and Pre-Test Questionnaire (see Appendix B), which provided details regarding their previous injury history, playing position and dominant foot, as well as confirming they were injury free and able to participate in the testing session. Anthropometric tests were then conducted to measure the height and body mass of each player. Each participant's height was measured while barefoot, to the nearest cm, using a Leicester Height Measure (Child Growth Foundation, UK). Body Mass was also recorded for each individual, to the nearest kg, using a set of electronic weighing scales (Seca, USA).

3.2.5 FUNCTIONAL MOVEMENT SCREEN SUBTESTS

The seven Functional Movement Screen subtests were then conducted on each player. The researcher conducted the FMS and is a Level 2 certified FMS tester. The order of the subtests and the scoring of the subtests followed the guidelines of Cook et al. (2006), to enable comparison with other research studies that followed the same protocol. This provided the researcher with a score out of 21 for each subject. In addition to the approved FMS scoring system, the researcher also followed a new customized scoring system for five of the seven tests, which offered each subject a score of between zero and four as outlined in Table 3.1. The existing FMS scoring can lead to subjects receiving the same total score despite vast differences in their performance on the subtests. The introduction of an increased range of scoring was aimed at improving the accuracy of the results for each player and potentially improving their correlation to injury prediction. Visual examples of the revised scoring system can be seen in Figure 3.1. Each participant completed each subtest three times, and for those that measure unilateral movement, three times on each leg or arm. The shoulder mobility, trunk stability push up and rotary stability tests required a clearing test to be completed to establish if any pain was present. If pain was reported, a score of zero was recorded for that test. In total, 105 players were tested on all seven of the FMS subtests and three clearing tests.

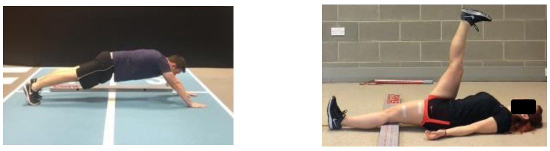
Table 3.1. The new scoring criteria applied for five of the seven FMS subtests

Test Name	Clearing Test	New Scoring Method Applied	New Scoring Criteria
Deep Squat	No	Yes	4 = As per normal 3 3 = As per 4 but with femur horizontal 2 = As per 4 but with feet on 6 x 2 board 1 = Unable to perform without compensation
Hurdle Step	No	No	N/A
Inline Lunge	No	No	N/A
Shoulder Mobility	Yes	Yes	4 = Hands within half a hands length 3 = As per normal 3 2 = As per normal 2 1 = As per normal 1
Active Straight Leg Raise	No	Yes	4 = Ankle joint moves past dowel residing in line with the Greater Trochanter 3 = As per normal 3 2 = As per normal 2 1 = As per normal 1
Trunk Stability Push Up	Yes	Yes	4 = Hands placed one hands length above top of head 3 = As per normal 3 2 = As per normal 2 1 = As per normal 1
Rotary Stability	Yes	Yes	4 = As per normal 3 3 = As per 4 but with up to 45-degree rotation of the trunk 2 = As per normal 2 1 = As per normal 1

3 on DS
Thighs Parallel

4 on SM
Fists within half
a hands length

3 on RS
Up to 45º rotation
of the trunk



4 on TPSU
Hands placed on hand
length above top of head

4 on ASLR
Medial malleolus resides
past the Greater trochanter

Figure 3.1. Visual examples of the new scoring criteria applied in five of the seven FMS tests

3.2.6 PERFORMANCE TESTS

To effectively screen a player for possible injury detection, the performance tests needed to mirror some of the actions performed during a soccer game. The research of Strudwick, Reilly & Doran (2002) and Reilly & Gilbourne (2003) was considered when prescribing the performance tests for this study. In addition, the type of injuries that soccer players suffered was also considered to assess potential injury risk. Research by Ekstrand et al. (2009) was used to understand the most frequent injury type for soccer players and to establish which tests would be most useful in identifying risk of such injuries. The five performance tests selected were as follows and were conducted in the order specified to ensure fatigue did not affect the performance of each player.

Single Leg Hop Test

This was performed twice per leg with a rest period of 30 s between each attempt. Participants placed the front of their take off foot on the start line. In their own time they jumped off one foot as far forward as possible, keeping their hands behind their back at all times. Their one-foot landing was then held for five seconds for the jump to count. The distance from the start line to the heel of their standing foot was then measured in cm, using a tape measure (Stanley, UK). The average distance jumped on each leg was calculated from the two jumps performed. The protocol used for this test was taken from research completed by Brummit et al. (2013) who found a link between single leg hop distance jumped as a percentage of height and lower back or lower extremity injury risk in male subjects.

Drop Jump/Vertical Jump Test

This was performed three times with a 30 s rest in between each jump. Participants stepped from a 30 cm high plyometric box landing on two feet in a squat position on to a jump mat placed 30 cm in front of the box, and then performed a vertical jump as quickly as possible, landing back on the jump mat (see Figure 3.2). For each jump the participant was asked to place their hands on their hips and to keep them in that position during test performance. The drop jump landing was recorded from the frontal plane using a Sony Handheld camera set at four and a half metres from the front of the plyometric box and analysed using Kinovea software. The knee-knee distance and ankle-ankle distance upon first contact with the mat during landing were measured and the absolute difference between the two was calculated. It was assumed that a perfect landing was one where there was no difference between the knee-knee distance and the ankle-ankle distance, and as such symmetry was achieved on landing. The average asymmetry distance for each player was calculated using all three landings. Vertical jump height was recorded using a Just Jump (Probotics, USA) mat and handheld recorder, giving a jump height in inches, which was then converted to cm. The average jump height for each player was calculated from the three jumps performed. The protocol used for this test was taken from research completed by Noyes (2005) but adapted to include measurement of ankle separation distance as well as knee separation distance in order to allow a comparison between the two upon landing.

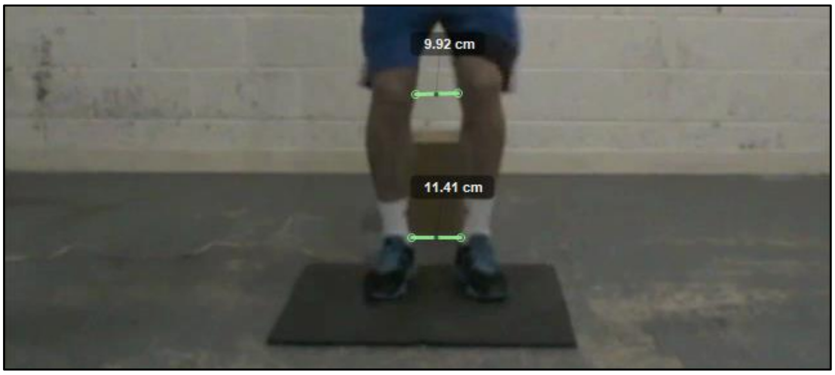


Figure 3.2. The Drop Jump Test. At first contact with mat, the knee-knee distance and ankle-ankle is measured, and the absolute difference between the two calculated

Lower Extremity Functional Test (LEFT)

This test was performed once. Participants undertook 16 shuttle runs around a diamond shaped course that challenged their agility, change of direction and speed. The distance between cones A and C, at the top and bottom of the diamond, was 9.14 m and cones B and D, in the middle of the diamond, was 3.05 m (see Figure 3.3). A Sony handheld camera was placed 11 metres from cone D, perpendicular from same to record each run. The overall time to complete the course was then measured using Kinovea software. The protocol used for this test was taken from research by Brummit et al. (2013) who reported a link between decreased LEFT time and increased risk of lower extremity injury or a foot or ankle injury in male subjects.

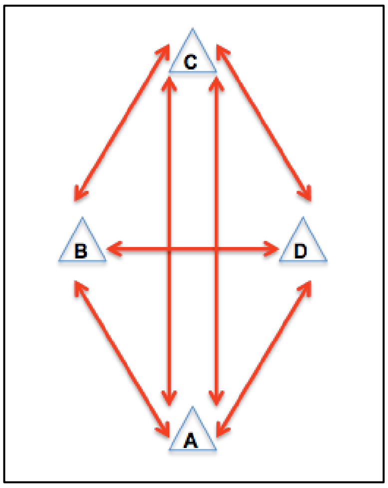


Figure 3.3. The LEFT. Distance between A and C is 9.14 metres, and B and D is 3.05 metres. Athlete completes a series of 16 shuttles as described in Appendix C

Yo-Yo Intermittent Recovery Test Level 2 (Yo-Yo)

This test was performed once at the end of the testing session and at least one hour of recovery time was given between this test and the previous test. All participants from each club undertook this test together. It required a space of 25 m by 20 m depending on numbers participating. Shuttle distance was 20 m, with an additional five metres allowed for recovery (see Figure 3.4). Performance was measured using the Yo-Yo Test (Top End Sports, UK) software and each player reached a certain stage and shuttle on the test, which was recorded on to the Screening Test Score Sheet (see Appendix D), along with the total distance in metres covered during the test. The protocol used for this test was taken from research by Bangsbo et al. (2008) which reported that the Yo-Yo Intermittent Recovery tests were a valid and simple tool by which to assess an individual's capacity to perform repeated intense exercise.

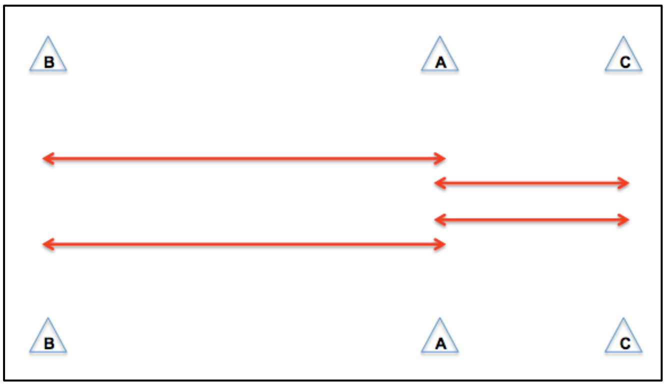


Figure 3.4. The Yo-Yo Test. Distance between A and B is 20 metres. Distance between A and C is 5 metres. Players start at A, run to B arriving on the beep, turn and run back to A arriving on the next beep. Players then have 10 seconds recovery to walk between A and C and back before repeating course

3.2.7 INJURY TRACKING

Following completion of the testing protocol, participants were tracked throughout the 2014 season regarding their injury occurrence rates. The club's physiotherapists were asked to record injury occurrence for each player during both training and match situations using the Injury Data Collection sheet (see Appendix E) provided by the researcher. The data recorded in the Injury Data Collection sheet included injury type, injury location, injury mechanism, number of days of absence and the severity of each injury, categorized as follows; Slight (1 to 3 days absence), Minor (4 to 7 days absence), Moderate (8 to 28 days absence) and Severe (more than 28 days absence). This is in line with the methods of The UEFA Injury model as described by Hägglund et al. (2005). The number of matches and training sessions each player was exposed too, were also recorded (see Appendix F). This data was sent electronically to the researcher once a month by the club physiotherapists for the researcher to place on the Injury Data Master sheet for all players. Injury records for 99 players were obtained throughout the season. Due to the manager of one of the clubs leaving that club within one month of the season starting, records for 16 players were not obtained and as such their data was removed from the results analysis. In addition, one player received a long-term illness within seven days of completing the testing protocol and was unable to continue his participation in the study. As such his data was removed from the results analysis. This left 99 players whose data was included in the final analysis.

3.2.8 STATISTICAL ANALYSIS

Using IBM SPSS software (version 21), the Shapiro Wilks test of normality was run on the data to establish how it was distributed and to determine the appropriate statistical test type. Frequency tests were performed on the injury data to establish the most common injury locations, types and mechanisms. Further analysis identified the percentage of injuries occurring during training and match situations and highlighted the most common timing for an injury to occur. The average length of absence per injury was also calculated, in addition to the level of training and match exposure for each player. For training exposure, the average training sessions per squad per season were identified to estimate the amount of training exposure per player. For match exposure, the incidence of injury per player per 1000 match minutes and non-contact injury incidence per 1000 match minutes per player were calculated (Equations 3.1 & 3.2). Descriptive statistics were performed to establish the mean total FMS composite score for each player using both the normal FMS scoring system and revised FMS scoring system. Spearman's Rank-Order Correlation tests were performed to identify if any correlations were present between players age, body mass or height and their FMS score or injury occurrence, using an alpha level of $p<0.05$.

Equation 3.1. Formula used to calculate incidence of all injuries per 1000 match minutes

$$(Number\ of\ injuries\ received \div total\ match\ minutes)\ \times 1000$$

Equation 3.2. Formula used to calculate incidence of non-contact injuries per 1000 match minutes

$$(Number\ of\ non-contact\ injuries\ received \div total\ match\ minutes) \times 1000$$

FMS and injury

An independent t test was performed to identify if significant differences existed in mean FMS scores and mean revised FMS scores between injured and non-injured players, between players with a non-contact injury and those not suffering a non–contact injury, and between players with an injury absence of eight days or more and players with an injury absence of seven days or less, using an alpha level of p<0.05. The same test was conducted to assess if significant differences existed in match injury incidence per 1000 match minutes between players scoring 14 or below and players scoring above 14 on the FMS tests. This test was repeated using a cut-off score of 16 or below on the FMS tests. The same test was performed for non-contact injury incidence per 1000 match minutes using both cut-off scores and repeated for the revised FMS scoring system using a cut-off score of 16.

Odds ratios and 95% confidence intervals were calculated, using 2x2 contingency tables, to establish the likelihood of various cut-off scores on the FMS increasing the likelihood of receiving an injury, non-contact injury and absences of eight days or more, as per previous research (Kiesel et al., 2007; Chorba et al., 2010). The sensitivity and specificity of the FMS cut-off scores were also calculated. The sensitivity was used to identify if a cut-off score of 14 or below on the FMS could accurately be used to identify players who received any type of injury (see Equation 3.3). The specificity was used to identify if a cut-off score of above 14 on the FMS could accurately be used to identify non-injured players (see Equation 3.4). Additional odds ratios were calculated for different cut-off scores for the FMS, for all types of injuries, non-contact injuries and absences of eight days or more. Odds ratios were calculated using a 2x2 contingency table for the revised FMS scoring system and different cut-off scores. The cut-off scores that maximised sensitivity and specificity was deemed to be that which produced the highest percentage when adding the sensitivity and specificity scores together. Chi square tests were run to establish if the odds ratio calculated for each cut-off was significant using an alpha level of p<0.05.

Equation 3.3. Formula used to calculate sensitivity

$$Sensitivity = \frac{True\ Positives}{True\ Positives + False\ Negatives} \times 100$$

Equation 3.4. Formula used to calculate specificity

$$Specificity = \frac{True\ Negatives}{True\ Negatives + False\ Positives} \times 100$$

Performance tests and injury

To understand if significant differences existed for performance test scores between injured or non-injured players, and between players not suffering non-contact injuries and players with non-contact injuries, Independent t tests were conducted with an alpha level of p<0.05. Using the mean times or distances achieved on each performance test as the cut-off scores, 2x2 contingency tables were produced to determine the odds ratio and 95% confidence interval levels of these cut-off scores increasing a player's likelihood of receiving any type of injury and a non-contact injury. Chi square tests were performed to determine the significance of the odds ratios calculated, and the sensitivity and specificity of each cut-off score were also recorded. An Independent t test was conducted to assess if significant differences existed in injury incidence per 1000 match minutes between players achieving the mean performance test score or below and players achieving above the mean performance test score. This was also determined for non-contact injury incidence per 1000 match minutes.

FMS scores and performance tests

Independent t tests were conducted to establish if significant differences existed in performance test scores between players with an FMS cut-off score of 14 or below and players with a score of above 14, using an alpha level of p<0.05. Spearman's Rank-Order Correlation tests were performed to identify if any correlations were present between total FMS composite score and a player's performance test score, using an alpha level of p<0.05, as well as between total FMS composite score using the revised scoring system and a player's performance test score.

3.3 RESULTS

3.3.1 PARTICIPANTS

Pre-season testing, and injury tracking, was conducted on 99 participants (mean±SD: age 23.2±4.4 years old; height 179.5±6.6 cm; body mass 77.5±7.8 kg). Seventy seven percent of all participants included in the results analysis were right foot dominant. Midfield was the most common playing position (30.2%), with centre forward (22.9%) and centre back (13.5%) the next two most common playing positions, across the study sample (Figure 3.5).

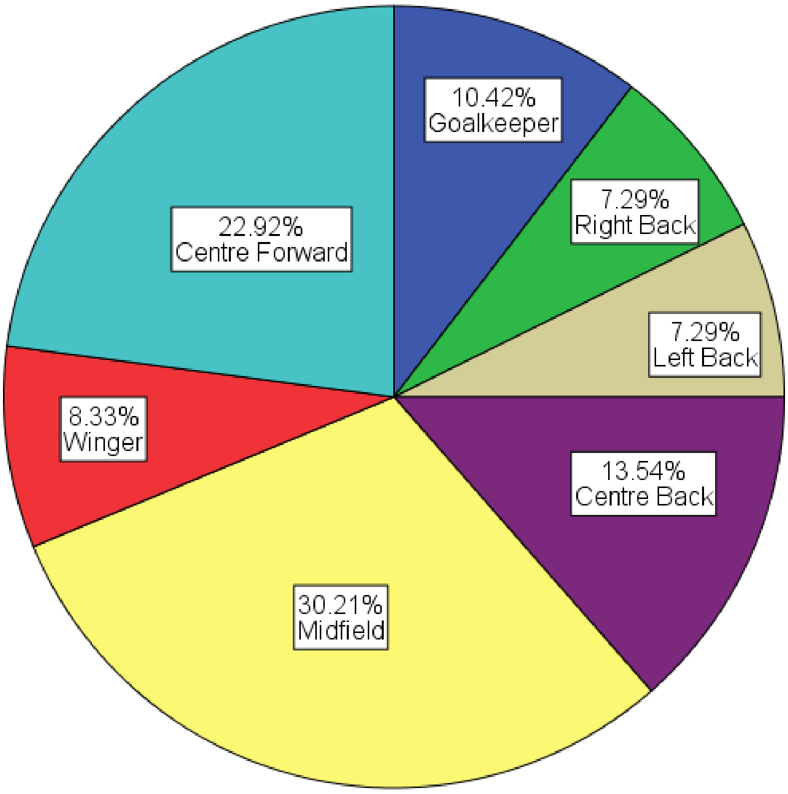


Figure 3.5. Percentage of different playing positions in sample population

The mean training minutes exposure across the five teams, throughout the 2014 season, was 8,874 minutes per squad, with a range of 6,920 to 9,990 minutes. It was not possible to identify the training exposure per player due to the data collection methods employed. The total match minutes for all 99 players were 136,423 minutes for the 2014 season. Mean match minutes per season per player were 1,378 minutes (Table 3.2).

Table 3.2. Mean match minutes per player per team for the 2014 season

Team	Total Match Minutes	Number of Players	Mean Match Mins per player
Team 1	24268	24	1011
Team 2	35157	20	1758
Team 3	35325	26	1359
Team 4	21454	15	1430
Team 5	20219	14	1444
Totals	136423	99	1378

3.3.2 INJURIES

In total, 109 injuries were recorded during the 2014 season. Non-contact injuries accounted for 60.6% of all injuries recorded, with 64.2% of all injuries occurring during matches. Of these, 80% occurred during the second half of matches. As per previous research (Ekstrand et al., 2009) the most common location of injuries reported were the thigh (31.6%) and ankle (21.1%). The average time loss absence per injury from soccer activity was 12.4 days and the average number of training/match days missed was 5.17 per injury. No significant correlation with injury occurrence was recorded for age (p=0.89), height (p=0.71) or body mass (p=0.48). For the sample population, match injury incidence was calculated to be 0.51 injuries per player per 1000 match minutes. When considering only non-contact injuries, match injury incidence was calculated to be 0.33 non-contact injuries per player per 1000 match minutes.

3.3.3 FMS AND INJURY RISK

Mean FMS total scores for the 89 players that undertook the seven FMS tests were 16.3 (See Table 3.3). No significant correlation with FMS scores was recorded for age (p=0.45), height (p=0.92), or body mass (p=0.86).

Table 3.3. Mean FMS scores for injured and non-injured players

Injured	N	Mean FMS Score	SD	Min	Max
Yes	57	16.3	2.0	10	19
No	32	16.3	1.5	14	20
Total	89	16.3	1.8	10	20

No significant difference was found between the mean FMS scores of injured and non-injured players when considering all injury types. Additionally, no significant difference was found in FMS scores between players not suffering a non-contact injury and players receiving a non-contact injury ($p>0.05$) (see Table 3.4). The mean FMS scores for injured (n=57) and non-injured players were 16.3 for both groups (SD±2.0 and 1.5). Forty players received a non-contact injury with a mean FMS score of 16.3 (SD±1.6).

Table 3.4. Mean FMS scores for players suffering a non-contact injury and players not suffering non-contact injuries

Non-contact Injury	**N**	**Mean FMS Score**	**SD**	**Min**	**Max**
Yes	40	16.3	1.6	13	19
No	49	16.3	1.9	10	20
Total	89	16.3	1.8	10	20

Following the procedures used in earlier studies (Kiesel et al., 2007; Chorba et al., 2010; Appel, 2012) 2x2 contingency tables were produced using various FMS cut-off scores to determine the odds ratio of a player receiving an injury based on their FMS score (see Appendix D.1 and D.2). The cut-off scores ranged from 13 to 18 and the odds ratio of receiving any type of injury, non-contact injuries and moderate/severe injuries (>8 days or more) were assessed for each cut-off. The odds ratio of receiving any type of injury ranged from 0.4 using a cut-off score of 17 to 1.28 using a cut-off score of 15. For non-contact injuries the odds ratios ranged from 0.39 using a cut-off score of 13 to 1.47 using a cut-off score of 16. Finally, for injuries lasting 8 days or more the odds ratios ranged from 0.76 to 1.89, using a cut-off score of 14. None of the cut-off scores reported for any injury type were significant ($p>0.05$).

The majority of previous research in this area has concentrated on a cut-off score of 14, with several studies reporting an increase in the odds of receiving an injury in participants scoring 14 or less on the FMS (Chorba et al., 2010; Duke et al., 2014; Kiesel et al., 2007; Garrison et al., 2015). In the current study, the odds ratio of a player scoring 14 or less on the FMS receiving any type of injury was 1.01 and was not significant, with a sensitivity of 16% and specificity of 84%. This odds ratio suggests a cut-off score of 14 has no better than a 50/50 chance of accurately predicting the odds of receiving all types of injuries.

When considering non-contact injuries, and using a cut-off score of 14, the odds ratio of these players receiving a non-contact injury was 0.63 and not significant, with a sensitivity of 13% and specificity of 82%. This odds ratio suggests that a cut-off score of 14 has a less than 50/50 chance of correctly predicting non-contact injury odds. When taking into account only moderate or severe injuries (>8 days), in line with previous research (Kiesel et al., 2007), using a cut-off score of 14, the odds ratio was 1.89 but not significant, and produced a sensitivity of 21% and specificity of 88%. This odds ratio suggests that a cut-off score has a slightly better than 50/50 chance of predicting the odds of receiving a moderate or severe injury.

When considering injury incidence levels, there was no significant difference for either all injuries or non-contact injuries received between those scoring 14 or below and those scoring above 14 on the FMS (All injuries, 0.51 vs 0.49 injuries per player per 1000 match minutes; Non-contact injuries, 0.36 vs 0.29 non-contact injuries per player per 1000 match minutes). Using a cut-off score of 16, which maximised sensitivity and specificity, there was no significant difference for either all injuries or non-contact injuries received between those scoring 16 or below and those scoring above 16 on the FMS (All injuries, 0.74 vs 0.67 injuries per player per 1000 match minutes; Non-contact injuries, 0.51 vs 0.38 non-contact injuries per player per 1000 match minutes).

Revised Scoring System

Using an Independent t test, no significant difference was found in mean FMS total composite score using the revised scoring system, as detailed in Section 3.2.5, between injured and non-injured players (p=0.89). Using the revised scoring system, the mean total FMS composite score for all participants who were tested on the FMS during pre-season (n=89), was 17.5 (SD±2.4). For injured players the mean score was 17.6 (SD±2.6) compared to 17.5 for non-injured players (SD±1.9) (see Table 3.5). For players with a non-contact injury the mean score on the revised FMS scoring system was 17.7 (SD±2.2) (see Table 3.6).

Table 3.5. Mean FMS scores using the revised scoring system for injured and non-injured players

Injured	N	Mean FMS Score	SD	Min	Max
Yes	57	17.7	2.0	10	22
No	32	17.6	1.9	14	21
Total	89	17.6	2.4	10	22

Table 3.6. Mean FMS scores using the revised scoring system for players with a non-contact injury and players not suffering a non-contact injury

Non-contact Injury	N	Mean FMS Score	SD	Min	Max
Yes	40	17.7	2.2	13	22
No	49	17.6	2.5	10	22
Total	89	17.6	2.4	10	22

Using a 2x2 contingency table, odds ratios were calculated to understand if a particular cut-off score increased the odds of a player receiving an injury using the revised FMS scoring system (see Appendix D.3 and D.4). The cut-off scores ranged from 14 to 20 and the odds ratio of receiving any type of injury, non-contact injuries and moderate/severe injuries (>8 days or more) were assessed for each cut-off. The odds ratio of receiving any type of injury ranged from 0.39 using a cut-off score of 19 to 1.5 using a cut-off score of 16. For non-contact injuries, the odds ratios ranged from 0.32 using a cut-off score of 14 to 1.85 using a cut-off score of 16. Finally, for injuries lasting 8 days or more the odds ratios ranged from 0.4 using a cut-off score of 20 to 2.2 using a cut-off score of 15. None of the cut-off scores reported for any injury type were significant ($p>0.05$).

Using the revised scoring system, a cut-off score of 16 maximised sensitivity and specificity for all injury types, non-contact injuries and absences of eight days, yet produced no significant odds ratios. For any type of injury, a score of 16 or less on the FMS using the revised scoring system had an odds ratio of 1.5 and produced a sensitivity of 33% and specificity of 75%, but this odds ratio was not significant. The same cut-off score produced an odds ratio of 1.85 of a player receiving a non-contact injury and produced a sensitivity of 38% and specificity of 76%, but once again was not significant. When only taking account of absences of eight days or more, a score of 16 gave a non-significant odds ratio of 1.98 and produced a sensitivity of 38% and specificity of 76%.

When considering injury incidence levels, there was no significant difference for either all injuries or non-contact injuries received between those scoring 16 or below and those scoring above 16 on the revised FMS scoring system (All injuries, 0.92 vs 0.60 injuries per player per 1000 match minutes; Non-contact injuries, 0.72 vs 0.32 non-contact injuries per player per 1000 match minutes).

Asymmetries

Thirty-four players had no asymmetries, 40 players had one asymmetry, 12 players were assessed to have two asymmetries and three players had three asymmetries as identified by the FMS. No significant difference was found in injury occurrence rates for all types of injury (p=0.42) or non-contact injuries (p=0.15) between player's with at least one asymmetry (n=55) on the FMS and players with no asymmetries (n=34). Using a 2x2 contingency table, the odds ratio of players with at least one asymmetry receiving any type of injury or a non-contact injury were not significant (see table 3.7). When considering injury incidence levels, there was no significant difference for either all injuries or non-contact injuries received between those with no asymmetries and those with one asymmetry or more (All injuries, 0.73 vs 0.68 injuries per player per 1000 match minutes; Non-contact injuries, 0.58 vs 0.36 non-contact injuries per player per 1000 match minutes).

Table 3.7. Summary of all injuries and non-contact injuries received by players with and without asymmetries

Injury Type	Asymmetry Yes	Asymmetry No	Odds Ratio	CI 95%	*p*	Sensitivity	Specificity
All Injuries							
Injured	37	20	1.44	0.59-3.49	0.42	67%	41%
Not Injured	18	14					
Non-Contact Injuries							
Injured	28	12	1.90	0.79-4.59	0.15	51%	65%
Not Injured	27	22					

Due to the small numbers with three or more asymmetries (n=3) it was decided not to analyse the difference in injury occurrence rates between this group and players with two or less asymmetries. No significant differences in injury occurrence rates for all injury types were found between players with two or more asymmetries (n=15) on the

FMS and players with one or less asymmetries (n=74). Using a 2x2 contingency table, the odds ratio of players with two or more asymmetries receiving any type of injury was not significant (see table 3.8). There was also no significant difference in injury incidence levels between players with two asymmetries or more and players with one or less asymmetries (0.59 vs 0.73 injuries per player per 1000 match minutes).

Table 3.8. Summary of all types of injuries received based on number of asymmetries

All Injuries							
Asymmetries (n)	**Injured (n)**	**Not Injured (n)**	**Odds Ratio**	**95% CI**	***p***	**Sensitivity (%)**	**Specificity (%)**
≥2	10	5	1.15	0.36-3.71	0.82	18%	84%
≤1	47	27					

No significant difference in non-contact injury occurrence was found between players with two or more asymmetries (n=15) on the FMS and players with one or less asymmetries (n=74). Using a 2x2 contingency table, the odds ratio of players with two or more asymmetries receiving a non-contact injury was not significant (see table 3.9). There was also no significant difference in non-contact injury incidence levels between players with two asymmetries or more and players with one or less asymmetries (0.34 vs 0.47 non-contact injuries per player per 1000 match minutes).

Table 3.9. Summary of non-contact injuries received based on number of asymmetries

Non-contact Injuries							
Asymmetries (n)	**Injured (n)**	**Not Injured (n)**	**Odds Ratio**	**95% CI**	***p***	**Sensitivity (%)**	**Specificity (%)**
≥2	7	8	1.09	0.36-3.31	0.88	18%	84%
≤1	33	41					

FMS subtests

Mean scores for the seven FMS subtests ranged from 2.92 on the trunk stability push up test to 1.96 for the hurdle step test. No significant difference was found in mean FMS subtest scores for any of the seven subtests between injured and non-injured players, between players suffering a non-contact injury and players not suffering a non-contact injury, or between players suffering a moderate or severe injury and those players who did not suffer a moderate or severe injury (p>0.05).

2 x 2 contingency tables were produced for each of the seven subtests, which analysed the odds ratio of an individual who scored a two or less on each subtest receiving any type of injury or a non-contact injury, compared to an individual who scored a three on each subtest (see Appendix D.5). The results showed that an individual scoring a two or less on the rotary stability subtest had a significant (p=0.05) odds ratio of 3.64 of receiving a non-contact injury, with a sensitivity of 47% and specificity of 83%. No other subtests produced a significant odds ratio of an individual receiving any type of injury or a non-contact injury based on scoring a two or less on any subtest (p>0.05). For all injuries, the odds ratios ranged from 0.54 for the inline lunge subtest to 2.84 for the rotary stability subtest. For non-contact injuries, the odds ratios ranged from 0.18 for the trunk stability subtests to the afore mentioned 3.64 for the rotary stability subtest. It should be noted that the small number of scores of three when performing the rotary stability subtest (n=5) resulted in a very small subgroup, which may could have influenced the results outlined above.

3.3.4 PERFORMANCE AND INJURY

To understand if a relationship existed between the performance tests and injury risk, statistical analysis was conducted on the performance test results. Mean test scores were calculated for injured and non-injured players, and for player's not suffering a non-contact injury and players suffering a non-contact injury, and then analysed for any significant difference. In addition, using the overall mean test scores as a cut-off point, odds ratios were calculated to ascertain if test performance increased the likelihood of receiving an injury (see Appendix D.6).

Vertical Jump Test

The mean height jumped was 51.7 cm (SD±5.3) for all participants who undertook the vertical jump test (n=72). No significant difference was found using an Independent t test in vertical jump height between injured (51.7 cm, SD±5.0) and non-injured players (51.8 cm, SD±6.0) (p=0.95). No significant difference was found in vertical jump height between players not suffering a non-contact injury (51.2 cm, SD±2.2) and those with a non-contact injury (52.3 cm, SD±1.9) (p=0.38). Using a 2x2 contingency table, the odds ratios of players who jumped the mean height or below (51.7 cm) receiving any type of injury (0.87) or a non-contact injury (0.52) were not significant. When considering injury incidence levels, there was no significant difference for either all injuries or non-contact injuries received between those with a vertical jump height of 51.7 cm or below and those with a height above 51.7 cm (All injuries, 0.41 vs 0.83 injuries per player per 1000 match minutes; Non-contact injuries, 0.31 vs 0.51 non-contact injuries per player per 1000 match minutes).

Lower Extremity Functional Test (LEFT)

The mean time recorded for all participants who completed the LEFT (n=27) was 110.54 s (SD±5.17). No significant difference was found using an Independent t test in mean LEFT time between injured (110.53 s, SD±4.8) and non-injured players (109.83 s, SD±5.9) (p=0.56). No significant difference was found in mean LEFT time between players not suffering a non-contact injury (110.11 s, SD±5.8) and those with a non-contact injury (111.17 s, SD±4.3) (p=0.60). Using a 2x2 contingency table, the odds ratios of players who recorded a mean time of 110.54 s or below receiving any type of injury (1.54) or a non-contact injury (1.2) were not significant. When considering injury incidence levels, there was no significant difference for either all injuries or non-contact injuries received between those with a LEFT time of 110.54 s or below and those with a LEFT time above 110.54 s (All injuries, 0.54 vs 0.36 injuries per player per 1000 match minutes; Non-contact injuries, 0.19 vs 0.19 non-contact injuries per player per 1000 match minutes).

Drop Jump Test

The mean drop jump asymmetry per player during landing was 2.6 cm (SD±2.0) for all participants who completed the drop jump test (n=67). No significant difference was found using an Independent t test in drop jump asymmetry on landing between injured (2.2 cm, SD±1.9) and non-injured players (2.6 cm, SD±2.1) (p=0.47). No significant difference was found in drop jump asymmetry on landing between players not suffering a non-contact injury (2.2 cm, SD±1.8) and players with a non-contact injury (2.7 cm, SD±2.3) (p=0.36). Using a 2x2 contingency table the odds ratios of players with a drop jump asymmetry of above 2.6 cm receiving any type of injury (0.91) or a non-contact injury (1.05) were not significant. When considering injury incidence levels, there was no significant difference for either all injuries or non-contact injuries received between those with a drop jump asymmetry score of 2.6 cm or below and those with an asymmetry score above 2.6 cm (All injuries, 0.38 vs 0.52 injuries per player per 1000 match minutes; Non-contact injuries, 0.25 vs 0.22 non-contact injuries per player per 1000 match minutes).

The Yo-Yo Intermittent Recovery (Yo-Yo) Test

The mean distance covered during the Yo-Yo test was 1376 m (SD±389) for all participants who participated in the test (n=53). No significant difference was found using an Independent t test, in the mean distance covered on the Yo-Yo test between injured (1388 m, SD±394) and non-injured players (1355 m, SD±403) (p=0.77). No significant difference was found in the mean distance covered on the Yo-Yo test between players not suffering a non-contact injury (1406 m, SD±380) and those with a non-contact injury (1342 m, SD±403) (p=0.55). Using a 2x2 contingency table, the odds ratios of players who recorded a mean distance of 1377 m or above receiving any type of injury (2.74) or a non-contact injury (1.97) were not significant. When considering injury incidence levels, there was no significant difference for either all injuries or non-contact injuries received between those recording a mean distance of 1377 m or above and those recording a distance of 1376 m or below (All injuries, 0.51 vs 0.78 injuries per player per 1000 match minutes; Non-contact injuries, 0.37 vs 0.47 non-contact injuries per player per 1000 match minutes).

Single Leg Hop Test (SLH)

On the single leg hop tests the mean distance jumped was 141.4 cm (SD±19.1) for the left leg and 141.9 cm (SD±18.7) for the right leg for all participants who performed the tests (n=50). No significant difference was found using an Independent t test, in SLH distance jumped between injured players (left leg 140 cm, SD±20.4; right leg 139 cm, SD±19.4) and non-injured players (left leg 143.6 cm, SD±17.3; right leg 146.7 cm, SD±16.9) (p=0.51 left leg; p=0.15 right leg). No significant difference was found in SLH distance jumped between players not suffering a non-contact injury (left leg 140.6 cm, SD±17.8; right leg 145.5 cm, SD±16.7) and players with a non-contact injury (left leg 142.3 cm, SD±21.1; right leg 137.8 cm, SD±20.3) (p=0.76 left leg; p=0.15 right leg). Using a 2x2 contingency table the odds ratios of players who jumped 75% of their height or below receiving either any type of injury (left leg 0.76; right leg 2.63) or a non-contact injury (left leg 0.67; right leg 2.59) were both not significant for either leg. When considering injury incidence levels, there was no significant difference for either all injuries or non-contact injuries received between those jumping a mean distance of 75% of their height or below and those jumping a distance of above 75% of their height (All injuries, 0.81 vs 0.42 (left leg) 0.56 vs 0.58 (right leg) injuries per player per 1000 match minutes; Non-contact injuries, 0.52 vs 0.33 (left leg) 0.23 vs 0.53 (right leg) non-contact injuries per player per 1000 match minutes).

3.3.5 FMS AND PERFORMANCE

Links have previously been established between performance and an individual's FMS score. The researcher attempted to explore if similar links existed within this sample population. Table 3.10 highlights the correlations found between FMS scores and performance tests scores using both the normal and revised FMS scoring system.

Table 3.10. Correlations between FMS scores (standard and new scoring system) and performance test scores

		Vertical Jump	**Drop Jump**	**LEFT**	**Yo-Yo**	**SLH Left**	**SLH Right**
Overall FMS Score	Correlation Coefficient	.244*	.156	-.326	.397**	.139	.051
	Sig. (2-tailed)	.044	.221	.104	.005	.340	.730
	N	68	63	26	48	49	49
Overall New FMS Score	Correlation Coefficient	.182	.255*	-.438*	.533**	.199	.103
	Sig. (2-tailed)	.138	.044	.025	.000	.171	.479
	N	68	63	26	48	49	49

*p<0.05 **p<0.01

Vertical Jump Test

Using an Independent t test, the difference in vertical jump height between players with a mean FMS score of 14 or less (46.9 cm, SD±3.2) and players with a mean FMS score of above 14 (52.4 cm, SD±5.3) was significant (p=0.006). When analysing correlation between FMS score and vertical jump height using a Spearman's Rank-Order Correlation test a significant, but weak, positive correlation was reported (r=0.24, p=0.04). No significant correlation was found between FMS score and vertical jump height when using the revised FMS scoring (r=0.18, p=0.13).

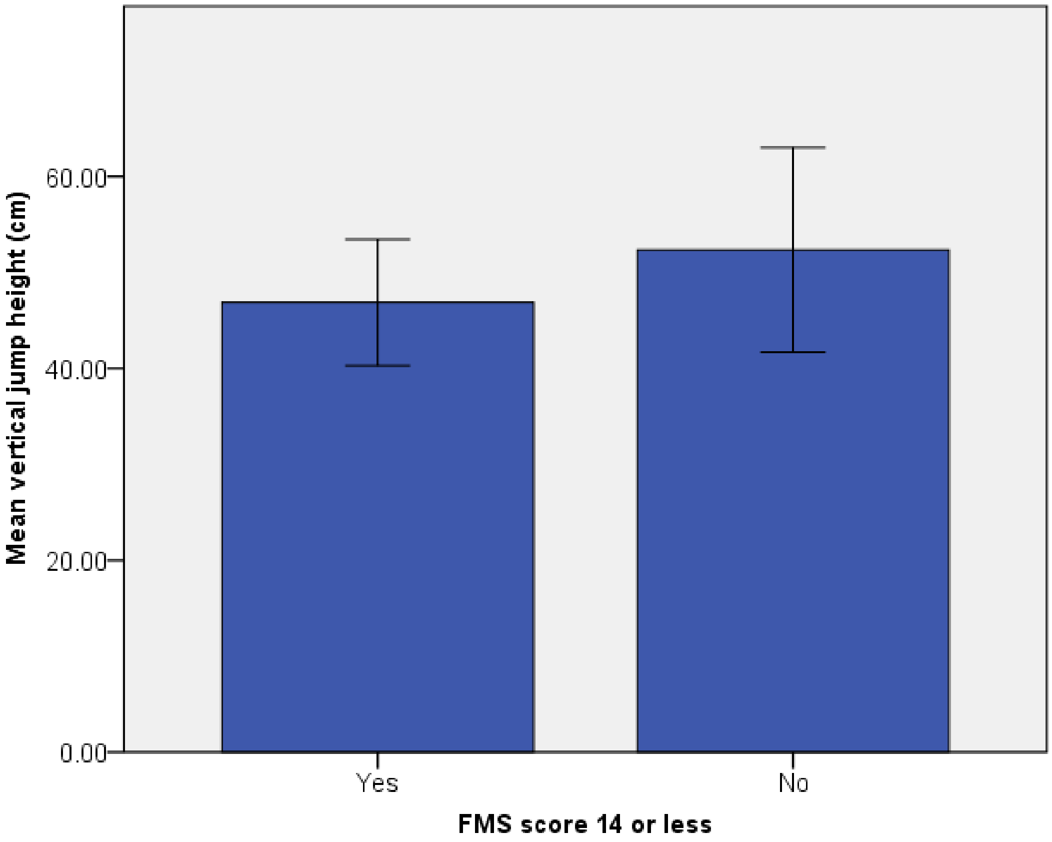


Figure 3.6. Difference in mean vertical jump height between players scoring 14 or less and above 14 on the FMS (difference is significant, $p<0.05$)

Lower Extremity Functional Test (LEFT)

Only two players who participated in the LEFT recorded a score of 14 or less on the FMS tests. Due to this number being extremely low, the significance of the difference in LEFT performance between those scoring 14 or less and those scoring above 14 was not calculated. A significant, moderate, negative correlation was found between FMS score and LEFT time ($r=-0.44$, $p=0.03$) but only when using the revised FMS scoring system.

Drop Jump Test

No significant difference was found in the mean distance in landing asymmetry distance on the drop jump between players with a mean FMS score of 14 or less (2.4 cm, SD±1.8) and players with a mean FMS score of above 14 (2.4 cm, SD±2.2) ($p=0.98$). A significant positive, but weak, correlation was reported between FMS score, using the revised scoring system, and drop jump landing asymmetry distance ($r=0.26$ $p=0.04$). This suggests that the greater the landing asymmetry distance on the drop jump the higher a player's FMS total score would be when using the revised FMS scoring system.

The Yo-Yo Intermittent Recovery (Yo-Yo) Test

No significant difference was reported in Yo-Yo test performance between players with a mean score of 14 or less on the FMS (1132 m, SD±398) and those players with a mean FMS score of above 14 (1387 m, SD±394) (p=0.17). A significant, moderate, positive correlation was found between FMS score and Yo-Yo test performance (r=0.40, p<0.01). Using the revised FMS scoring system, a stronger correlation was reported (r=0.53, p<0.01).

Single Leg Hop Test

No significant difference was reported in single leg hop test distance in either leg between players with a mean FMS score of 14 or less (136.4 cm, SD±12.4 left leg; 140.6 cm, SD±13 right leg) and those players with a mean FMS score of above 14 (141.4 cm, SD=19.7 left leg; 141.6 cm, SD±19.4 right leg) (p=0.51 left leg; p=0.89 right leg). No significant correlation was found between FMS scores and single leg hop distance in either the left leg (r=0.14) or the right leg (r=0.05). Using the revised FMS scoring system, no significant correlation was reported for either the left leg (r=0.20) or the right leg (r=0.10).

3.4 DISCUSSION

Main Findings

Previous research has indicated that an FMS cut-off score of 14 or less increases the odds of a player receiving a moderate or severe length injury (Kiesel et al., 2007). The current research study found that using this same cut-off score, there was no significant increase in a player's odds of receiving any type of injury, a non-contact injury or moderate/severe injury. Additionally, no significant difference was found in injury incidence levels for all injuries or non-contact injuries received between player's scoring 14 or less on the FMS and those scoring above 14. When considering individual subtest scores, the rotary stability subtest produced a significant (p=0.05) odds ratio of 3.64 of a player who scored two or less receiving a non-contact injury. Although this subtest provided a positive and significant odds ratio, the number of individuals in the sample who scored a three on this subtest (n=5) means that this subgroup is too small to report that the rotary stability subtest can accurately be used to predict a player’s odds of receiving a non-contact injury. Further research using a sample with a wider spread of scores on the rotary stability subtest would be useful to understand if a link does exist between subtest score and non-contact injury risk. Additional subtests did not produce any significant odds ratios of a player scoring two or less on any subtest receiving any type of injury or a non-contact injury, suggesting subtest score cannot accurately be used to predict an individual’s odds of receiving an injury.

When considering performance tests and injury risk, no significant difference was found in any of the five performance test scores between injured and non-injured players, or between players not suffering a non-contact injury and those with a non-contact injury. The results indicated that Yo-Yo test performance was negatively correlated with injury risk, and although this trend was non-significant, it suggests that those performing better on the Yo-Yo tests are at an increased risk of injury.

The current study also analysed the relationships between FMS scores and performance test scores and found that a score of 14 or less on the FMS was linked with a significantly lower performance on the vertical jump test. Other performance tests did not report any significant difference in performance using the same cut-off score. Significant positive correlations were reported regarding a players FMS score and their performance on the vertical jump test and Yo-Yo test.

Injuries

The most common injury type and locations reported in the current study are in line with that of previous research (Woods et al., 2002; Ekstrand et al., 2009). Muscle strains, ligament sprains and contusions were the most frequently occurring injury types, and the thigh and the ankle were the most common injury locations amongst this sample population. This suggests there is little difference between professional and semi-professional soccer players in regard to their level of risk of receiving a particular injury type in a certain location. These findings are surprising given that professional players experience significantly greater levels of training exposure over one season than amateur players and would also have access to advanced levels of medical support and advice. Match injury occurrence in the current study accounted for a larger percentage of all injuries compared to training injury occurrence than in earlier research (Woods et al., 2002), However as per previous studies, match injuries were significantly more common than training injuries, suggesting players are at an increased risk to injury during matches compared to training, possibly due to the intensity and amount of contact during matches being higher than in training situations. Practitioners should consider how to structure training to best prepare players for the rigours of matches, without increasing their injury risk during training sessions. When considering match injury incidence levels, the current study produced a similar incidence level of injuries per player per match minutes/hours as previous studies (Soderman et al., 2001; Ekstrand et al., 2009) despite the exposure levels recorded being significantly lower than that in other studies, and the playing level to be lower. The timing of match injuries in the current study is also consistent, with previous research, with a greater percentage of injuries occurring in the second half of matches, and in particular the last 15 minutes of the second half. The role fatigue plays in second half injury occurrence cannot be under estimated and suggests the use of monitoring techniques to track individual players during matches is an area that still needs to improve if match injury rates are to reduce.

FMS and injury

A non-significant odds ratio of 1.89 was reported for players scoring 14 or less on the FMS receiving a moderate length injury. Despite this higher odds ratio for moderate length injuries compared to those found for all injuries (1.01) and non-contact injuries (0.63), the results do not support the previous research in this area that found a substantial odds ratio (11.67) of athletes scoring 14 or less receiving a moderate length injury (Kiesel et al., 2007). When adjusting the cut-off score to one that optimised sensitivity and specificity (a cut-off score of 16) the odds ratios produced for all injury types were still not significant but were similar to that seen in previous studies (Appel, 2012; Letafatkar et al., 2014). The odds ratios produced were similar for all injury types (1.02 vs 1.01) higher for non-contact injuries (1.47 vs 0.63) but reduced for moderate length injuries or longer (1.89 vs 1.61). These results suggest that the FMS cannot accurately be used to predict a male soccer player’s odds of receiving any type of injury and using only the total FMS composite score within the current study was not a reliable indicator of a soccer player’s injury risk.

The impact of FMS score on non-contact injury risk was analysed in the current study. The FMS is used to measure an individual's fundamental movement patterns, and as such it was assumed FMS score would be more closely linked with non-contact injury risk than all types of injury. Previous research in this area (Kiesel et al., 2007; Peate et al., 2007; Chorba et al., 2010) made no reference to injury type in their studies and as such only considered links between FMS score and certain injury types or severities. The current research did not produce any stronger odds ratios of a player receiving a non-contact injury based on FMS score, compared to the odds ratios produced for all injury types. The low number of non-contact injuries (n=40) could have impacted these findings and reduced the statistical power of the tests undertaken.

The current study recorded match exposure levels to assess if including number of match minutes played increased the odds of receiving an injury, when taking into account FMS score. As stated in earlier paragraphs, no significant difference was found in injury incidence levels for all injuries or non-contact injuries received between player's scoring 14 or less on the FMS and those scoring above 14 (All injuries, 0.51 vs 0.49 injuries per player per 1000 match minutes; Non-contact injuries, 0.49 vs 0.36 non-contact injuries per player per 1000 match minutes). These findings differ from previous research that has found a link between a score of 14 or less and increased injury risk (Kiesel et al., 2007; Chorba et al., 2010). Including exposure levels helps to normalise injury data, so essentially provides a more accurate record of injury occurrence rates and non-contact injury occurrence rates. Earlier research studies that have reported a link between FMS score and injury (Kiesel et al., 2007; Letafatkar et al., 2014) did not include any analysis of exposure levels, and as such make no allowance for the impact of training or match minutes on injury risk, which could be of more relevance than their FMS composite score in regard to injury occurrence. The fact that the results from the current study suggest that a player's FMS score does not significantly increase their injury incidence or non-contact injury incidence rate per 1000 match minutes, would indicate that total FMS composite score cannot be used as the only predictor of injury within applied settings.

It has been suggested that the current FMS scoring system may not be the most effective way to measure the FMS subtests (Gulgin & Hoogenboom, 2014; Parenteau et al., 2014). The introduction of the revised FMS scoring system within this current study attempted to increase the accuracy of the FMS through clearer assessment of each subtest performance. However, this revised system did not produce any significant odds ratios, using various cut-off scores, of a player receiving any type of injury and the range of odds ratios produced were similar to those seen using the existing FMS scoring system. It would therefore seem that the revised scoring system is no more accurate than the existing FMS scoring system in predicting injury risk within a population of soccer players. Due to the extra time constraints required to conduct the revised scoring system as it is currently designed, it would not be of benefit to employ this revised system in an applied setting. If further work in this area were to be considered it should perhaps focus on a major revamp of the FMS scoring system rather than the small adjustments to the scoring system employed in this study.

The FMS allows a practitioner to identify players with asymmetries on five of the seven subtests that they complete. It is suggested that those with asymmetries make compensations in their fundamental movement patterns (Cook et al., 2006), and as such one could assume, would be more at risk of injury than those with no asymmetries. Players within this sample population that had at least one asymmetry on the FMS had no significant increase in their injury incidence compared to those players with no asymmetries. The same result was also found when considering only non-contact injuries. In addition, no significant odds ratios were reported for players with one or more asymmetry, in relation to their chances of receiving any type of injury or a non-contact injury. The results from the current study suggest that analysing male soccer player’s asymmetries will not provide any indication of their injury risk, and as such are no more relevant than using their FMS composite score, as per previous research (Chorba et al., 2010; O’Connor et al., 2011), to predict their risk of injury. There was no significant increase in injury incidence levels per player per 1000 match minutes for all injuries or non-contact injuries received between players with two or more asymmetries and those with one asymmetry or less. The small sub group size could be a factor in these results as only 12 players had two asymmetries or more, although they appear to indicate that using the number of asymmetries to predict a player’s injury risk is not reliable.

Previous research has reported a link between lower scores on the deep squat and trunk stability subtests and non-contact injury (Rusling et al., 2015). The current study did not find the same relationship for these two subtests but did find that a score of two or less did produce a significant odds ratio of 3.64 of receiving a non-contact injury. Whilst the sub group of players scoring a three on the rotary stability subtest was low (n=5), the findings suggest that further research in this area could be worthwhile. Whilst there is insufficient evidence from the current study, or indeed previous research, to suggest that subtest scores can accurately predict non-contact injury, most of the research in this area to date has not even considered if a relationship between subtest scores and injury exists (Chorba et al., 2010; Kiesel et al., 2007; Letafatkar et al., 2014). Given that each subtest is designed to test either mobility, stability or full body patterning, hence placing the body in challenging positions that replicate everyday activities, it would seem reasonable to suggest a weakness, dysfunction or asymmetry in one particular area, leading to a low score on a particular subtest, could be linked to an increase in the odds of receiving an injury. In the case of the rotary stability subtest, multi-plane, pelvis, core and shoulder girdle stability are assessed, and the movement requires proper neuromuscular co-ordination and energy transfer in the torso. Poor performance on this subtest can be attributed to poor reflex stabilisation or the trunk and core, as well as reduced hip stability and limited mobility in the knee, hip, spine and shoulder. Therefore, focussing on the individual subtest scores may help to provide a more accurate reflection of the ability of the FMS to predict injury risk, as it involves an individual’s weakest link in regard to movement competency being considered as an injury risk, removing the chance for weaker scores to be hidden by good performances on other subtests through the use of the total FMS composite score, which could be negatively impacting the results.

Performance tests and injury

No significant difference was found in any of the five performance test scores between injured and non-injured players, or between players not suffering a non-contact injury and those with a non-contact injury. The Yo-Yo test produced an odds ratio of 2.74 for players achieving a mean distance of 1377 m or more receiving any type of injury and an odds ratio of 1.97 of receiving a non-contact injury. These odds ratios were not statistically significant, but the findings suggest that in this group of players there was a negative trend between Yo-Yo performance and injury risk. One reason for this trend could be that players with a greater level of fitness are likely to play more matches over a season, and therefore would seem to be at increased risk of injury, when in fact their risk is no higher than those with low Yo-Yo scores e.g. if a player with a high Yo-Yo score plays 1000 match minutes and is injured twice in that period, compared to a less fit player who only plays 500 minutes but is injured once, they actually have the same level of injury risk per 1000 match minutes. This is highlighted, when considering exposure levels in the current study, where no significant difference in injury incidence occurrence per player per 1000 match minutes was found based on Yo-Yo test performance.

Research by Brummit et al. (2013) has suggested the LEFT was reliable at predicting lower extremity injury risk amongst 193 college athletes. They reported that male athletes completing the LEFT in 100 s or less were 3.2 times more likely to receive a time loss injury to the lower back, and 6.7 times more likely to receive a foot or ankle injury, than those completing the LEFT in 101 s or more. In the current study, the LEFT produced a non-significant odds ratio of greater than one of receiving all injuries or a non-contact injury, but these were significantly lower than those recorded in the previous research. These results indicate that the Yo-Yo test and LEFT performance cannot be used to accurately identify a player’s risk of injury. The Single Leg Hop test was the only other performance test to provide an odds ratio greater than one of receiving an injury, but this was also not significant. Players jumping less than 75% of their height on their right leg had an odds ratio of 2.63 of receiving any type of injury and an odds ratio of 2.59 of receiving a non-contact injury. However, the same relationship was not found on the left leg.

FMS scores and performance tests

A score of 14 or less on the FMS tests is linked with a significantly lower performance on the vertical jump test. Significant positive correlations were reported regarding a players FMS score and their performance on the vertical jump test and Yo-Yo test. A previous relationship between the countermovement jump and FMS scores has been reported by Conlon (2013), in his small study of 36 male students. He established that countermovement jump height was significantly correlated with FMS scores using a live 21-point scoring scale ($r=0.34$, $p=0.03$) and a video 21-point scoring scale ($r=0.43$, $p<0.01$). These findings are similar to those from the current study and suggest further work in this area, is worthy of consideration. Further research could use a prospective, randomised controlled study with a larger sample size, where FMS scores are increased in an intervention group and not the control group, to assess if improvements in jump height are also found in the intervention group. When using the revised FMS scoring system, a significant positive correlation was found between a player’s revised FMS score and performance on the Yo-Yo test, and a significant negative correlation was found between a player’s revised FMS score and performance on the LEFT and Drop Jump test. Waldron et al. (2014) suggested the FMS was not effective in measuring a person’s athletic ability, and Parchman & Mcbride (2011) suggested the FMS is not an adequate field test to determine athletic performance. The results from the current study would further indicate that the FMS cannot be used accurately to determine performance, but that some relationship may exist between FMS scores and performance in the vertical jump test and FMS score and Yo-Yo test performance. In addition, the results suggest a relationship may exist between revised FMS scores and LEFT performance and revised FMS scores and drop jump performance. The correlations that exist are not strong enough to state that FMS scores could accurately predict each test performance with detailed precision but do suggest a lower FMS score might indicate a worse performance on the vertical jump test and Yo-Yo tests, and a higher revised FMS score might indicate a better performance on the LEFT and a greater landing asymmetry on the drop jump test.

3.5 STUDY STRENGTHS

3.5.1 INJURY DEFINITION

Having a clear definition of injury is important for any injury surveillance study and should help to provide the reader with clear and accurate results that can be applied in the field. This study used a time loss definition of injury, in line with the UEFA model (Hagglund et al., 2005), which meant that any injury that resulted in an absence from participation in a training session or match, that occurred during a club training session or match, was recorded and included in the results analysis. This has ensured that the injury recording procedures could be consistent across all players. It has also allowed for a more complete analysis of the data to be completed, and in particular to assess any links between FMS scores and all types of injury, something that has been missing from previous studies in this area. In addition, this study has considered non-contact injuries and their links with FMS score. It was expected that including non-contact injuries would strengthen any links that might be seen between FMS score and injury prediction and provide greater insight into the role the FMS could play in injury prediction. However, these links were not seen within the current study.

3.5.2 EXPOSURE LEVELS

A further strength of this study is the inclusion of exposure levels in regard to match minutes for each player. From an analysis perspective this has meant that the links between FMS scores/performance test scores and injury can be more accurately analysed by normalising injury rates to the amount of injury exposure, which varies from player to player based on match exposure. Previous research has focussed on the FMS cut-off score of 14 or less and the number of injuries associated with individuals within this group (Chorba et al., 2010; Letafatkar, Hadadnezhad, & Shojaedin, 2014). No allowances for individual's exposure rates have been included within the analysis, which could have led to inconsistencies within the results. For example, players with increased match exposure levels will be at greater risk of injury than those players with limited match playing time. This increased risk is not because of their FMS score, it is simply due to their increased playing time. To date research studies have not discussed the effect this could have on the links identified between FMS score and injury risk.

3.5.3 REVISED FMS SCORING SYSTEM

When scoring the FMS, it is very common for two participants to receive a score of two on a particular subtest, despite there being obvious differences in the quality of their performance. A score of two is given to an individual who can complete the movement required in a subtest with some level of compensation. The FMS scoring system provides a broad range of acceptable criteria within which performances should fall to score a two on any subtest. This results in performances that are close to being scored a one, and other performances that are close to scoring a three, both being scored as a two. The revised scoring system employed was designed to allow individuals who would normally perform better on each subtest, but not necessarily see this reflected in the score provided using the existing FMS scoring system, to score higher compared to those individuals who would normally perform towards the lower end of the scoring continuum. It was anticipated that this would more accurately analyse each participant's mobility and/or stability on each subtest. In one of the subtests, the revised FMS scoring system broke down the mid-level performance into two scores instead of one (a score of two was split into a two or a three). In four of the subtests, the revised scoring system added an additional scoring level (a score of four was given for performance that would normally be classed as a high three).

3.5.4 INCLUSION OF PERFORMANCE TESTS

A final strength of this study is the inclusion of performance tests and their links with injury. Previous research has typically focussed on the links between FMS scores and injury risk, or FMS scores and their links with performance. No study to date, has compared the links between FMS scores and injury and performance test scores and injury, to see if any relationship exists between the two. This study has undertaken such a comparison, which has provided data that can be applied in the field when a practitioner is deciding on their own injury prevention/screening strategy.

3.6 STUDY LIMITATIONS

3.6.1 LENGTH OF STUDY

It was the intention of the researcher to complete a two-season study on this sample population. Due to the access provided by the clubs and players involved in the study, it was not possible to continue the study beyond one season. An extra season of injury tracking, along with an additional testing session with every participant, would have given the researcher further data that may have provided more insight into the links between FMS and injury.

3.6.2 SAMPLE SIZE

A major limitation of the current study was the sample size recruited. Although the sample size used for this study was comparable to other studies using sporting populations, and in many cases, was larger, a greater uptake from the clubs asked to participate in the study would have provided the study with greater statistical power. This particularly applies to carrying out sub group analysis in areas such as non-contact injuries and asymmetries, by providing a wider range of FMS scores, and additional injury data to allow more insightful analysis of the links between FMS tests, performance tests and injury. Once analysis moved from the overall group into smaller sub sections, the participant numbers became low for certain sub groups, reducing the statistical power in these sub group analyses, making it difficult to draw strong conclusions from some of the results.

Effect sizes were calculated (see Appendix D.7) for FMS and different types of injuries, FMS subtest scores and injury, and performance tests and injury, to determine the sample size required to establish if a real significant difference existed between groups. These calculations indicated that a sample size of 326 would be required for a study analysing if FMS score is significantly different between participants suffering an injury of eight days or more, and those not suffering such an injury. When considering FMS subtest scores, only two of the subtests had an effect size of 0.2 or above. The calculations suggest a sample of 253 for the inline lunge subtest and 155 for the rotary stability subtest, would be required to understand if subtest scores were significantly different between injured and non-injured participants. For the performance tests, only the LEFT had an effect size of 0.2 or above. The results indicated that a sample of 298 would be required to establish if LEFT time was significantly different between injured and non-injured participants. These results suggest that due to the very small effect sizes shown for several of the tests, sample sizes of approximately 150 to 300 would be required to test if these differences were actually statistically significant.

3.6.3 TRAINING EXPOSURE

Training exposure for each player was calculated with the assistance of the medical staff that were employed by the five clubs involved in the study. The nominated medical personnel were asked to record the number of participants at each training session throughout the season, and to record the length and date of each session. Due to the part time nature of the League of Ireland, the details recorded by each member of staff varied from club to club, and as such it was not possible to accurately assess the training exposure for each individual player. Although the training exposure per squad was accurately identified, the researcher acknowledges that the complete recording of training exposure for each player would have benefitted the study as it would have allowed for a more detailed analysis of injury incidence in training, and any links with FMS scores to be established.

3.6.4 INJURY TRACKING

The researcher followed the guidelines set out in the UEFA model (Hagglund et al., 2005) regarding tracking injury occurrence throughout the 2014 season. The methods employed were done so successfully and accurately. However, the researcher acknowledges that this element of the study remained outside of his control, as the onus was on the medical staff from the clubs involved to complete the Injury Data Collection form in an accurate and timely manner. Two clubs changed their medical staff during the study, and as such variability in the way injuries were recorded could have existed between different personnel. To reduce the possibility of this, clear definitions and guidelines were supplied by the researcher.

3.7 CONCLUSION

FMS and injury

This study suggests that the link between FMS score and injury risk in male soccer players is weaker than previously reported in other population groups, and that a cut-off score of 14 or 16 cannot predict, with any certainty, a player's injury risk. The link between FMS score and absences of eight days or more is stronger compared to other injury types in this study and could possibly reach statistical significance with greater participant numbers or by tracking injuries over a longer period. A link between FMS scores and injuries of 21 days or more (Kiesel et al., 2007) has been established in a different sporting population. Additionally, this study reported a significant increased odds ratio for players who scored low on the rotary stability subtest, receiving a non-contact injury. This suggests that future research in the area of FMS and injury should consider the inclusion of subtest scores in their injury analysis to fully understand if a link does exist between subtest scores and injury risk.

Performance tests and injury

The results suggest that the performance tests used in this study cannot accurately be used to predict a player's likelihood of receiving any type of injury or a non-contact injury, particularly when match exposure levels are considered. Although an individual's performance on the Yo-Yo test, LEFT and Single Leg Hop tests produced a higher odds ratio of receiving all types of injuries than a cut-off score of 14 on the FMS, the results are not strong enough to state that these performance tests are more effective in predicting injury than FMS scores.

FMS and performance tests

Significant correlations were reported between FMS scores and vertical jump, LEFT, Drop Jump and Yo-Yo test performance. This indicates that a relationship exists between different fitness variables and FMS scores in male soccer players. Although the correlations reported are not strong enough to suggest using FMS scores as the only predictor of performance, it does indicate that fundamental movement patterns do play a role in an individual's physical performance and improving these may improve performance test results.

Implications

These results suggest that the FMS cannot accurately predict a soccer player’s odds of receiving any type of injury. Previous research has shown a link in different sample populations, between moderate length injuries and longer, and a cut-off score of 14 on the FMS, but the current study did not show a significant increase in moderate length injury risk for player’s scoring 14 or below. As the current study has already identified, it is difficult to establish a single factor that causes an injury, or a single process or test that will help prevent the injury from occurring, and therefore using the FMS as a standalone screen to predict injury does not appear to be a valid practical use of the FMS.

Perhaps the main limitation of research to date in this area is using only the mean composite FMS total score as the primary indicator of injury risk. From an applied perspective, practitioners are unlikely to use this score as a tool to prescribe future interventions or indeed to predict injury risk, as the FMS manual (Functional Movement Systems, 2017) clearly specifies that individual subtest scores should be used to make decisions relating to future testing or interventions. In addition, as discussed in Chapter Two, it is difficult to accurately select a suitable cut-off score as it should depend on the type of sample used. Certain samples will have a higher level of physical competency than others, and as such a higher cut-off score would be required to effectively assess if the total FMS composite score can accurately predict an individual’s odds of receiving an injury. Future studies should consider if the using the total FMS composite score is the most effective method by which to assess injury odds, and if so, the most relevant cut-off score to use based on the physical competency levels of the sample used.

Future research should also consider including an intervention designed to improve FMS score, when assessing if the FMS is linked with injury prediction. To date, all studies that have assessed FMS and injury risk have simply measured FMS score across their sample, and then collected injury data on each participant, before analysing if there is a significant link between the two. Results to date using this approach have been inconclusive and as such, future studies using a prospective, randomised controlled research design where FMS score is increased in an intervention group and not in a controlled group, and injury rates tracked using a clear injury definition and including normalised injury data, could help to further knowledge in this area by reducing the effects of individual differences on injury rates. Such interventions would also be based on the individual subtest scores, hence ensuring that these studies take subtest scoring, and exposure levels, into consideration when assessing injury risk, rather than simply using a simplistic model of applying a cut-off score of 14.

The FMS does appear to have merit in regard to understanding an individual's fundamental movement patterns, and as such it is widely used and recommended for use by practitioners internationally. Coupled with the body of research that has found links between FMS scores and injury in different sample populations, there appears to be validity in exploring the possibility of updating or amending the current scoring system, which may improve the screens ability to predict injury and performance. As the testing procedure for the FMS is still relatively new, reducing the subjectivity seen in the scoring system, could help to improve the reliability of the screen. It has been shown that the reliability of the FMS reduces significantly when considering subtest scores (ref), yet these scores are of more relevance to practitioners when prescribing corrective exercise interventions or when deciding on the next course of action for each individual. The simplistic nature of the scoring system can lead to errors being made when scoring each subtest which lead to intrarater and interrater errors. In order to reduce the subjectivity of FMS scoring, two different approaches could be adopted. Firstly, the scoring system itself could be updated to include more detailed biomechanical measurements that provide objective data based on the application of certain measurement tools or devices, utilised alongside the FMS testing kit. Whilst this would provide more definitive measurements following test performance, introducing such a scoring system would not only increase the time required to conduct and analyse the FMS, but it would also fundamentally change the concept of the FMS as a simple, time efficient, screen that can be conducted with minimal equipment and in any environment. Therefore, any significant adjustment to the scoring system in an attempt to reduce the subjectivity of subtest scoring, is unlikely to be widely adopted by practitioners and as such would be of limited use. Perhaps a more effective approach to reducing subtest scoring subjectivity is to remove the manual nature of FMS scoring, through the development of an automated approach. Although a significant challenge, this approach would remove the manual aspect of the scoring process and make the scoring system more objective. This should lead to an increase in the reliability and accuracy of FMS subtest scoring, which will provide practitioners with more confidence in its results, hence increasing its use in applied settings. It would also offer researchers the opportunity to fully analyse the effect of both total FMS composite scores and subtest scores on injury risk and performance, using a valid and standardised measurement tool, reducing the chance of human error and bias effecting the results.

CHAPTER FOUR

DEVELOPMENT OF MOTION TRACKING SOFTWARE TO ASSESS THE DEEP SQUAT FMS SUBTEST

4.0 INTRODUCTION

As outlined in the previous chapters, research undertaken to date has reported links between the FMS and injury (Kiesel et al., 2007; Chorba et al., 2010; Letafatkar et al., 2014) and the FMS and performance (Conlon, 2013; Chapman et al., 2014). Some of these studies have reported strong links between an individual's FMS score and their potential risk of injury, and in particular their risk of receiving a moderate length injury or longer (eight days or more) (Kiesel et al., 2007). However, these research studies have included several limitations that bring into question the results produced, and the research study outlined in chapter three supports the view that FMS scores cannot accurately be used to predict injury risk in a population of soccer players.

The limitations of previous FMS research include the use of different injury definitions, lack of reference to match or training exposure levels and small sample sizes. Such sample size limitations could have arisen because of the time required to undertake a FMS assessment. This can result in the screening protocol becoming too intrusive for the teams and individuals involved and is particularly apparent when considering non-professional teams, where time is extremely limited and hence screening of athletes is not always seen as a priority. This can make it difficult to conduct studies lasting longer than one season and as such reduces the power of these studies. The ability to retest participants across concurrent seasons would allow further data to be collected and to assess the effects of any intervention programme employed, which to date has not been possible.

Additionally, inter-rater and test-retest reliability of the FMS is weaker when considering the subtests rather than just the total FMS composite score (Shultz et al., 2013; Gulgin & Hoogenboom., 2014). Research has identified that the rotary stability and inline lunge subtests in particular, are prone to lower levels of agreement between different raters (Minick et al., 2010; Shultz et al., 2013). The current scoring system is subjective and can lead to errors in the scores provided for each subtest which could impact any intervention programme prescribed. From an applied perspective, a practitioner would be keen to monitor the effectiveness of any intervention over a set period of time, which could involve reassessment of an individual's FMS performance several times over this period. If the test-retest reliability of the FMS, when considering subtest scores, is not high, it makes it difficult for a practitioner to understand if their interventions are effective or not. If the subjectivity of the scoring system could be reduced or removed, small changes in performance could be tracked on an ongoing basis, allowing the effectiveness of the intervention programme to be more accurately assessed, and if required, amended, based upon the results.

The FMS Manual (Functional Movement Systems, 2017) provides a detailed algorithm in relation to how to interpret subtest scores to effectively guide a practitioner in deciding on the correct intervention for each individual. A FMS rater should follow this process when assessing the FMS results, which focusses first on the mobility related subtests (active straight leg raise and shoulder mobility), then on the stability related subtests (rotary stability and trunk stability), before considering the remaining subtests that look at whole body movement. A score of one or an asymmetry on the left or right side of the body on any subtest, highlight areas of weakness, with priority given to low scores on the mobility subtests first, then the stability subtests and finally the whole-body movement subtests. By following this process, a practitioner can easily identify the areas to include in any intervention programme. In addition, this process ensures that the individual can focus on improving their weakest areas, and that they are kept away from placing their bodies in potential harmful positions that could cause long term injury. Improving the accuracy of subtest scores would help to improve the quality of any intervention programme prescribed and reduce the chances of individuals following programmes that have been prescribed based on incorrect subtest assessment.

The previous paragraph highlights why subtest scoring is the critical component of the FMS, and how the total composite score out of 21 is not particularly important in relation to programme prescription. Therefore, if the FMS is not reliable in correctly identifying subtest scores across multiple testers the usefulness of it as a screen in an applied setting is questionable. In addition, the validity of research discussing the FMS will continue to be heavily reliant on the level of experience of the testers involved until such subjectivity is removed. This makes critical evaluation of research and application of any results in the field very difficult.

The aim of this research study was, therefore, to develop novel automated software, linked to the Microsoft Kinect sensor, that could accurately track the deep squat FMS subtest and automatically score deep squat performance in line with standard FMS guidelines as specified in the FMS Manual (Functional Movement Systems, 2017). The deep squat subtest was chosen as it represented a total body movement pattern that involved multiple joints working in combination, starting from a bilateral standing position that can be assessed from the frontal or sagittal plane. It therefore represented a suitable challenge in regard to the tracking of key joint centres during a multi-joint movement that would provide the researcher with a good indication of the software's ability to track additional FMS subtests that could be incorporated into future work.

4.1 RESEARCH AIMS (PROTOTYPE V1)

This study attempted to answer the following research question; Could software, integrated with the Kinect sensor v1, be developed that can validly assess performance on the deep squat FMS subtest compared to manual scoring by a certified FMS tester?

4.2 METHODS (PROTOTYPE V1 DEVELOPMENT)

4.2.1 SELECTION OF MOTION CAPTURE TECHNOLOGY

When selecting the motion capture technology to integrate with the software development, it was important to take into consideration certain factors that would influence the practicality of using such software in an applied setting. As outlined in Chapter 2, several alternatives were available, including marker based motion capture systems, inertial measurement units and markerless depth cameras. Marker based motion capture systems (MBS) typically employ the use of small markers attached to the body of the subject and a set of two or more cameras focussed on the subject to capture its motions (Kolahi et al., 2007). They are often used to assess biomechanics relating to various sporting activities and medical diagnostics (Wang et al., 2003). Inertial measurement units (IMU's) are self-contained devices that provide dynamic motion information through direct measurements (Altun et al., 2010). They often contain either gyroscopes that provide angular rate information around an axis of sensitivity, or accelerometers that provide linear or angular velocity rate information, or feature both pieces of technology. Markerless depth cameras provide real time 3D video capture that can identify the human skeleton and joint centres without the need for markers to be placed on the participant. One such markerless system is the Microsoft Kinect sensor which was developed to support the Microsoft Xbox games console to provide players with the opportunity to play games without the need for handheld controllers due to the Kinect’s integrated 3D camera capabilities (Muller et al., 2017).

Each of these technologies had their strengths and weaknesses, so three key criteria were employed to make the decision regarding which technology to use. Firstly, the technology primarily needed to be used in an applied setting, whereby the hardware and software were portable and available to be used in multiple locations. Secondly, it was important that the technology could be set up and calibrated in an efficient manner without causing delays to testing sessions that involved large sample sizes. Finally, the ability to accurately use the technology repeatedly to ensure consistency in data capture was an important factor to consider when choosing the right technology to employ. Figure 4.1 outlines which of the technologies met these three criteria using a simple Yes/No system. This quickly provided a clear understanding of the most suitable technology to use for the development of new software to assess the deep squat FMS subtest.

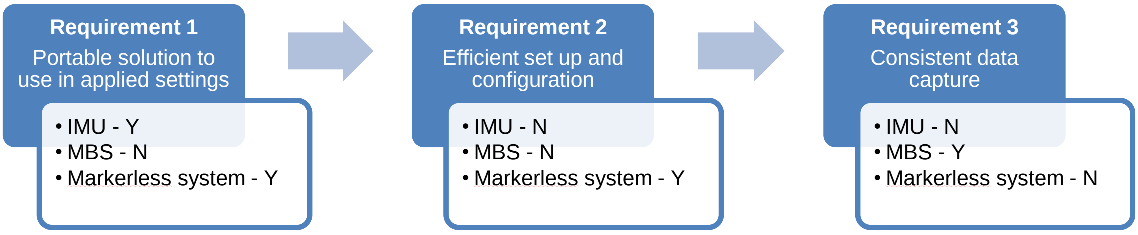


Figure 4.1. Overview of criteria used to select technology. IMU - Inertial measurement unit, MBS - Marker based system

Based on the criteria outlined in Figure 4.1 the markerless Kinect sensor was the most practical and cost-effective technology to integrate with bespoke software to assess the FMS deep squat subtest. Although this still presented challenges from a reliability and accuracy perspective, it offered a greater opportunity to be used in an applied setting, due to its ease of set up and calibration, non-invasive nature and accessibility/portability. As highlighted in Chapter 2, marker based systems are only reliable when the marker placement is accurate, which is unfortunately prone to human error and can be time consuming (Mündermann et al., 2006). The markerless nature of the Kinect sensor removes the variability of marker placement and therefore offered an opportunity for reliable bespoke software to be designed that was specifically tailored to the specific needs of FMS assessment.

4.2.2 TECHNOLOGICAL CAPABILITY OF MICROSOFT KINECT v1

The Kinect sensor v1 (Figure 4.2) was a markerless depth camera that houses an RGB camera along with an infrared depth sensor creating a 3D map allowing an individual's movements to be tracked. It consists of a horizontal bar with a small base that attaches to a laptop through a 3.0 USB connection. For the purposes of this study the Kinect sat in front of the laptop on a table 0.75m high, to provide an unrestricted view of the individual performing the movements. The Kinect v1 used infrared patterns to acquire depth information by reflecting these patterns off the person being tracked. It had a range of 1.2 to 3.5m, an angular field of view of 57º horizontally and 43º vertically and had a motorised pivot that could tilt the sensor up and down as required.

Figure 4.2. The Microsoft Kinect sensor Version 1 (Microsoft, 2015)

4.2.3 PROTOTYPE V1 DEVELOPMENT PHASE

To ascertain if software could be developed and integrated with the Kinect sensor v1, to automatically score and analyse the deep squat FMS subtest, a prototype model (v1) of the software was developed, that allowed the technology to be tested in an applied setting. This prototype was developed by the researcher in conjunction with the Telecommunications Software and System Group (TSSG), a software development team who are associated with Waterford Institute of Technology (WIT). The researcher led the project and used his expertise within the area of FMS to provide the specific requirements for the prototype to the TSSG development team. In turn, the development team focussed on writing the software code and algorithms to support these requirements. The researcher was also responsible for ongoing testing of the prototype to provide feedback to the development team to enhance development work. Using the Microsoft Kinect Software Developer Kit (SDK), the Microsoft Kinect sensor v1 was integrated with a Windows based laptop (Sony Viao) to create prototype v1. Once this integration was complete, code and algorithms were written to allow prototype v1 to track the movements of an individual completing the deep squat FMS test.

The primary aim of the software development was to ensure that each participant's movements, when completing the deep squat subtest, were captured by prototype v1 and sampled in real time through the laptop, to allow automatic and accurate assessment of the subtest. The scoring of the movement was based on the FMS scoring system provided in the FMS Manual (Functional Movement Systems, 2017), hence providing four possible scoring options. A score of three was given for a perfect performance of the deep squat as defined by the FMS Manual. A score of two was given for a perfect performance of the deep squat but with the participant standing on the 6ft by 2ft FMS board. A score of one was given if the individual was unable to perform the deep squat perfectly whilst using the FMS board. For the purposes of prototype v1 there was no option to score the movement as a zero if participants felt pain during the performance. Instead, the participant was asked to stop the performance if they were in pain.

Skeletal Tracking

Using the inbuilt infrared emitter, the Kinect sensor v1 recognised an individual within its field of view without the need for markers. To be recognised by the sensor an individual was required to stand facing the sensor between 1.2m and 3.5m away. The infrared emitter of the Kinect sensor v1 projected a pattern of infrared light which was used to calculate the depth of the person in the field of view allowing the recognition of different body parts and joints. A tracked skeleton provided detailed information relating to the individual in the field of view and recognised up to 20 joints on an individual's body as outlined in Figure 4.3.

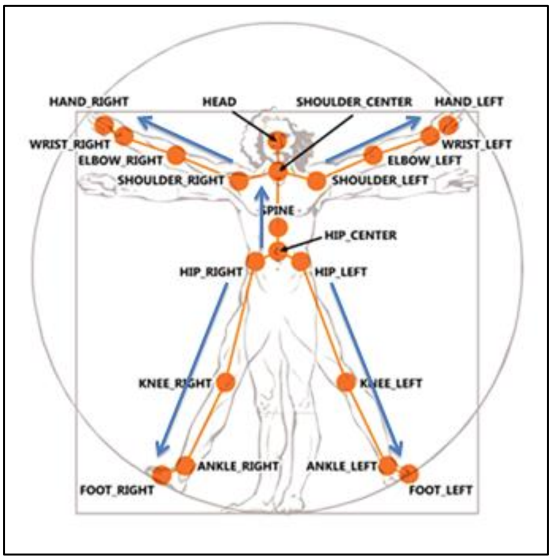


Figure 4.3. The 20 joints recognised by the Microsoft Kinect sensors Skeletal Tracking system (Microsoft, 2015)

Joint Orientation System

The Kinect SDK provided joint orientation information for the skeleton identified by the Skeletal Tracking system. A hierarchy of bones was defined using the joints identified by the Skeletal Tracking system. This hierarchy had the Hip Center joint as the root and extends to the feet, head, and hands (Figure 4.4). Bones were specified by the parent and child joints that enclose the bone. For example, the Hip Left bone was enclosed by the Hip Center joint (parent) and the Hip Left joint (child) (Figure 4.5). Hierarchical rotation provided the amount of rotation in 3D space from the parent bone to the child. This information outlined how much rotation in 3D space was needed in the direction of the bone relative to the parent. The rotation of the Hip Center joint provided the absolute orientation of the individual in camera space co-ordinates. This assumed that an individual has the origin at the Hip Center joint, the y-axis is upright, the x-axis is to the left, and the z-axis faces the camera.

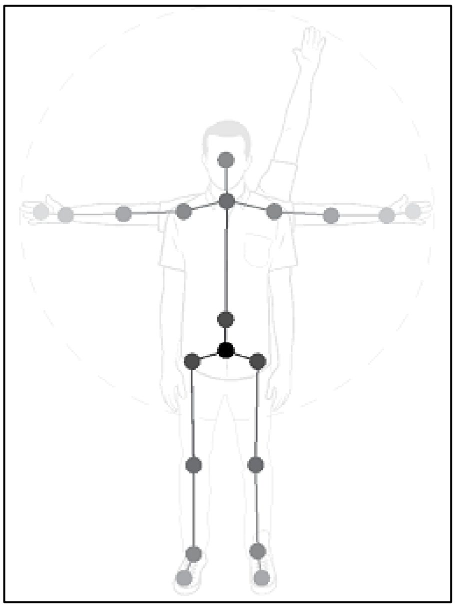


Figure 4.4. The Joint Orientation System with the Hip Center joint as the root (Microsoft, 2015)

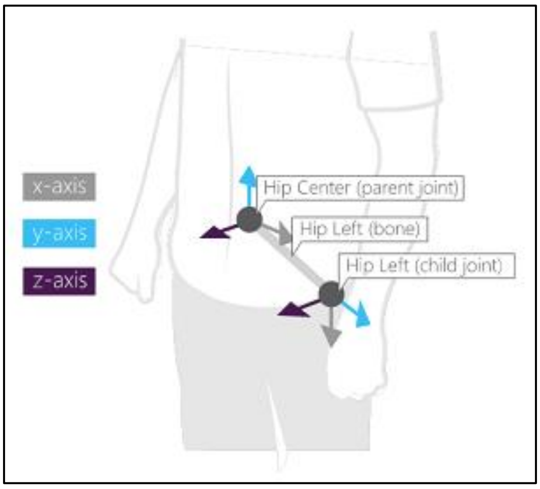
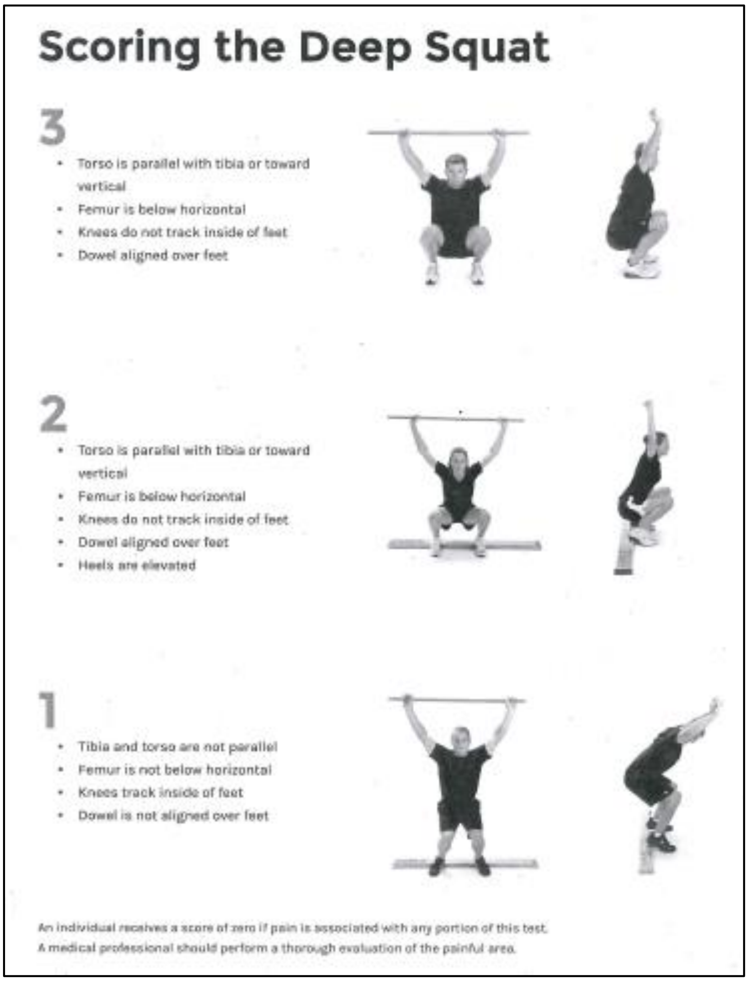


Figure 4.5. Hip Left bone is enclosed by the Hip Center joint (parent) and the Hip Left joint (child) (Microsoft, 2015)

Custom Gestures

The FMS Manual provides a description of the acceptable movement thresholds allowed within various body parts during the deep squat subtest and how these relate to a score of zero, one, two or three (Figure 4.6).



Scoring the Deep Squat

3

- Torso is parallel with tibia or toward vertical
- Femur is below horizontal
- Knees do not track inside of feet
- Dowel aligned over feet

2

- Torso is parallel with tibia or toward vertical
- Femur is below horizontal
- Knees do not track inside of feet
- Dowel aligned over feet
- Heels are elevated

1

- Tibia and torso are not parallel
- Femur is not below horizontal
- Knees track inside of feet
- Dowel is not aligned over feet

An individual receives a score of zero if pain is associated with any portion of this test.
A medical professional should perform a thorough evaluation of the painful area.

Figure 4.6. The deep squat FMS subtest scoring criteria (Functional Movement Systems, 2017)

As highlighted in Figure 4.6 there are four main scoring criteria that should be considered when scoring deep squat performance:

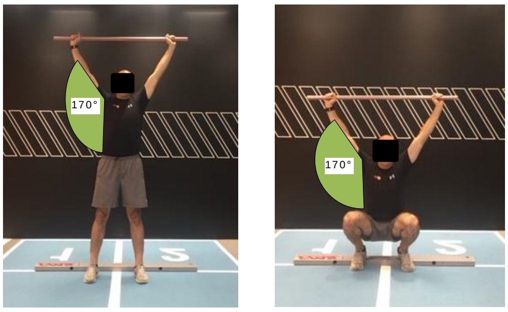
1. Position of torso in relation to tibia
2. Depth of the femur at bottom of squat
3. Position of knees in relation to feet
4. Position of dowel in relation to feet

These criteria are the same for scores of two and three, with the only difference being that the heels are raised on the FMS board for a score of two. If the scoring criteria are not achieved, then a score of one is given. Any feelings of pain during subtest performance results in a score of zero.

The joints specified in the Skeletal Tracking system were used to set custom gestures to assess deep squat performance using the Joint Orientation system to track movement. These gestures were written into the code supporting the Kinect SDK. The gestures are outlined in Tables 4.1 to 4.3, with visual examples provided in Figures 4.7 to 4.9. A pre-test start position gesture was set outlining the angle range deemed acceptable between the Wrist Right/Left and Ankle Right/Left in the frontal plane, using the Shoulder Right/Left as the vertex. During subtest performance, the angles between key joints were analysed to ascertain a score of one, two or three. This included measuring the angle between the Shoulder Right/Left and Ankle Right/Left using the Hip Right/Left as the vertex, the Hip Right/Left and Ankle Right/Left, using the Knee Right/Left as the vertex, and the Wrist Right/Left and Ankle Right/Left using the Shoulder Right/Left as the vertex.

Table 4.1. Custom Gestures used with the Shoulder Right/Left as the vertex for the deep squat subtest

	Score		
Shoulder Joint	**3**	**2**	**1**
Abduction/ adduction	In the frontal plane, angle with the Shoulder Right/Left as the vertex formed by two lines (one from Wrist Right/Left to Shoulder Right/Left and one from Ankle Right/Left to Shoulder right/Left) increases/decreases by less than 10 degrees during movement	As per 3 but standing on 6x2 board	In the frontal plane, angle with the Shoulder Right/Left as the vertex formed by two lines (one from Wrist Right/Left to Shoulder Right/Left and one from Ankle Right/Left to Shoulder right/Left) increases/decreases by 10 degrees or more during movement



[redacted]
[redacted]

A

B

Figure 4.7. Examples A and B highlight angles used to measure shoulder joint abduction/adduction in the frontal plane

Table 4.2. Custom Gestures used with the Hip Right/Left as the vertex for the deep squat subtest

	Score		
Hip Joint	**3**	**2**	**1**
Adduction/ abduction	In the frontal plane, angle with the Hip Right/Left as the vertex formed by two lines (one from Shoulder Right/Left to Hip Right/Left and one from Ankle Right/Left to Hip Right/Left) increases/decreases by less than 15 degrees during movement	As per 3 but standing on 6x2 board	In the frontal plane angle with the Hip Right/Left as the vertex formed by two lines (one from Shoulder Right/Left to Hip Right/Left and one from Ankle Right/Left to Hip Right/Left) increases/decreases by 15 degrees or more during movement

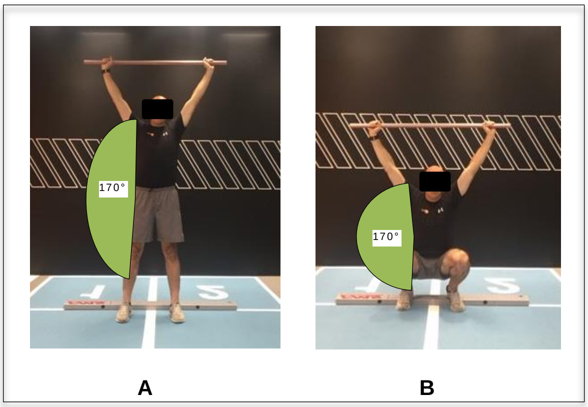


Figure 4.8. Examples A and B highlight angles used to measure hip joint abduction/adduction in the frontal plane

Table 4.3. Customer Gestures used with Knee Right/Left as the vertex for the deep squat subtest

	Score		
Knee Joint	**3**	**2**	**1**
Flexion	Angle with the Knee Right/Left as the vertex formed by two lines (one from Hip Right/Left to Knee Right/Left and one from Ankle Right/Left to Knee Right/Left)	As per 3 but standing on 6x2 board	Angle with the Knee Right/Left as the vertex formed by two lines (one from Hip Right/Left to Knee Right/Left and one from Ankle Right/Left to Knee Right/Left)

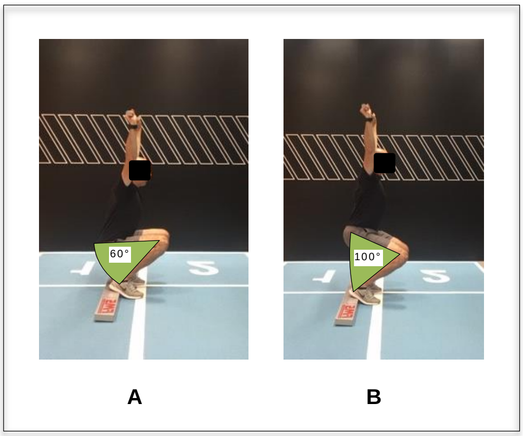


Figure 4.9. Examples A and B highlight angles used to measure knee joint flexion/extension in the sagittal plane

Software Operation

The operator began testing by pressing the Start Test button on screen (Figure 4.10), and the software indicated at this point if the individual was in the correct start position. The operator also had the ability to specify that an individual was standing on the FMS board, and hence prototype v1 recognised that a two was the maximum score available. Once individuals performed the deep squat movement and were at the base of the squat, the operator pressed the Finish Test button on screen. Prototype v1 automatically provided a score based on the performance at this point.

Figure 4.10. Screenshot from prototype v1

4.3 METHODS (PROTOTYPE V1 VALIDATION)

4.3.1 SAMPLE SIZE

Development of prototype v1 took eight weeks. Following completion of the development phase, validation testing was scheduled to test the validity of prototype v1 when assessing deep squat performance. A convenience sample of 141 school children (mean age=9.7years; SD±3.7) were screened on the FMS deep squat subtest as part of the validation study. This convenience sample was chosen as it allowed a wide range of movement styles and body sizes to be sampled in a short period of time. If prototype v1 could validly assess different movement styles and body sizes, then the researcher could be confident in the ability of the fully developed software to assess all types of body shapes and sizes, and different movement patterns. The validation study was completed in conjunction with the Castlepoint and Rochford School Sports Partnership in the UK and involved visiting seven schools over a week-long period in November 2014.

4.3.2 TESTING

A Sony Viao laptop connected to the Kinect sensor v1 was set up in each school hall and a cross marked on the floor two metres away from the Kinect sensor v1. The sensor and laptop were placed on a table set at 0.75 m high. Each class were divided into groups of five and one group at a time were tested. A member of school staff was present during every test and each participant was wearing P.E kit that included T shirt or tracksuit top and shorts/tracksuit bottoms with suitable trainers. A full demonstration of the deep squat subtest was provided by the researcher. Each participant was instructed to stand on the cross marked out on the floor and to face the Kinect sensor v1. One at a time each participant undertook the deep squat subtest three times whilst standing in front of the Kinect sensor v1. The researcher began each test by pressing the start button on the laptop screen, before instructing the individual to complete the deep squat movement. When the individual was in the finished position at the base of the squat the researcher pressed the finish button on the laptop screen, and recorded the best score provided by prototype v1 as well as the angles identified between key joints on the Prototype Testing Score Sheet (see Appendix D). In addition, the researcher manually scored the best subtest performance for each participant, as the performance was taking place, and recorded this score upon test completion, as soon as this was physically possible. Only the best score identified by prototype v1, and manually by the researcher, were recorded on the Prototype Testing Score Sheet due to time constraints. Although the researcher provided his own manual score for each performance, that was not based on the score provided automatically by prototype v1, the researcher was not blinded to prototype v1 scores, which is a limitation of this part of the study.

4.3.3 STATISTICAL ANALYSIS

To ascertain if prototype v1 could accurately assess the FMS deep squat subtest compared to an expert tester, statistical tests were performed using IBM SPSS software version 21. Mean deep squat scores across all participants were calculated for both prototype v1 and manual scoring methods. To assess the validity of prototype v1 as a method to assess deep squat performance, the absolute and percentage agreement between both scoring methods for all participants scores were calculated. Additionally, a Spearman rank-order correlation test was conducted to understand the level of correlation between prototype v1 and manual scoring, as another method to test the validity of the prototype. To analyse inter-rater reliability between prototype v1 and manual scoring, a weighted kappa (K_w) statistic was calculated. A weighted kappa score of 0.80 and above indicates excellent agreement, 0.60 to 0.79 represents substantial agreement, 0.40 to 0.59, moderate agreement and below 0.40, poor to fair agreement. This analysis is in line with previous research (Teyhen et al, 2012; Shultz et al, 2013; Parenteau et al, 2014).

4.4 RESULTS (PROTOTYPE V1)

In total, 141 participants performed the deep squat subtest three times. One participant reported feelings of pain when performing the subtest and was therefore asked to stop the test and their results were removed from the analysis. Mean deep squat subtest score was 1.3 (SD±0.5) for manual scoring and 1.2 (SD±0.6) for prototype v1 scoring. Absolute agreement between manual and prototype v1 scoring was perfect for scores of three and excellent for scores of two and one (Table 4.4). Overall agreement for all participants was 95%. Scores of two (n=39) and three (n=3) reported 100% agreement between the scoring methods, with all participants that were identified by manual scoring as a two or a three, also picked up automatically by prototype v1 as the same score (Figure 4.11). Seven individuals scored as a one by manual scoring were not given any score by prototype v1 as it failed to track their performance during the tests. When considering concurrent validity, there was a very strong positive correlation between the manual scoring and prototype v1 scoring methods $r_s = 0.99$ ($p<0.01$).

Table 4.4. Breakdown of absolute agreement between prototype v1 and manual scoring methods for scores of one, two and three on the deep squat FMS subtest. Figures in red highlight disagreement between scoring methods

	Manual Scores				
Software Scores	**1**	**2**	**3**	**No Score**	**Total**
1	91	0	0	0	91
2	0	39	0	0	39
3	0	0	3	0	3
No Score	7	0	0	0	7
Total	98	39	3	0	140

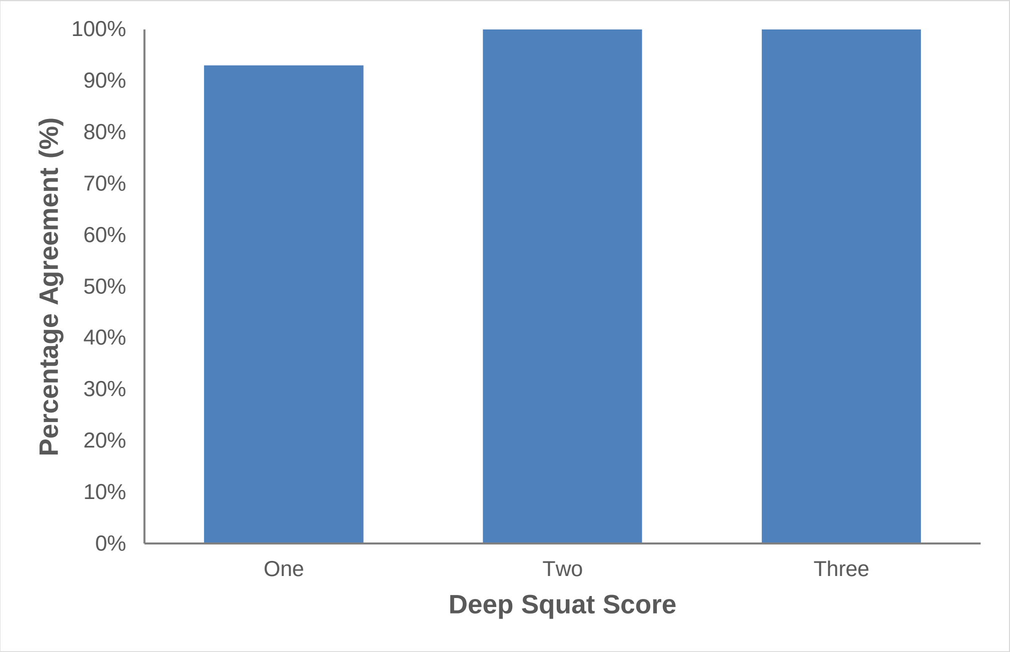


Figure 4.11. Summary of percentage agreements between prototype v1 and manual scoring methods for the deep squat FMS subtest (n=140)

From a reliability perspective, prototype v1 showed excellent reliability compared to manual scoring with statistically significant agreement between the two scoring methods (K_w=0.89).

4.5 DISCUSSION (PROTOTYPE V1)

Prototype v1 correctly identified 133 out of 140 participants performing the deep squat subtest as the same score as that provided by the certified FMS tester. The overall level of agreement between the scoring systems (95%) was higher than previous research that assessed agreement for the deep squat subtest between manual raters (Minick et al., 2010; Teyhen et al., 2012). In addition, a strong positive correlation was reported between the two scoring methods (r_s=0.99) suggesting that prototype v1 had potential to be used as a valid method for assessing the deep squat subtest. The reliability of prototype v1 when compared to manual assessment, using a weighted Kappa, showed excellent levels of reliability (K_w=0.89). This level of reliability is higher than previously reported inter-rater reliability levels between manual raters for the deep squat subtest (Minick et al., 2010; Teyhen et al., 2012).

Prototype v1 failed to recognise 7 of 140 participant’s performance as a deep squat, even though their performance was manually scored as a one by the certified tester. Prototype v1 was unable to track these participant's movements during the deep squat subtest through its Joint Orientation system and as such was unable to provide any score for the performance. This was despite prototype v1 correctly identifying the key joints of the body on each of these individuals using its Skeletal Tracking system. This highlighted a limitation of prototype v1 for the assessment of additional FMS subtests. Although an individual’s joints can be correctly identified pre-test, during the performance of a subtest involving multi-joint movement, prototype v1 was unable to accurately track these joints for all individuals. This issue was unexpected, as prototype v1 incorporated simple custom gestures and analysed only three key joints during subtest performance. This suggested that more custom gestures, using both joint angles and alignment could be beneficial to increase the accuracy of joint tracking during performance of the deep squat subtest, and indeed the remaining six FMS subtests.

It should also be noted that due to the high level of ones scored across the participants, the data was skewed and thus it was more likely that agreement between scoring methods would be high. Not only is a score of one the easiest of the three scores to identify as there are clear criteria that are not met when scoring a one, but also one possible score (a three) is already ruled out due to the participant’s heels being placed on the FMS board. This results in a 50/50 chance of the participant scoring a one before they have even begun performing the subtest, increasing the odds of agreement between manual and prototype v1 scoring significantly. Due to the age of the participants, the quality of the performances on the deep squat subtest was low, with many participants lacking the necessary co-ordination and motor control required to have a chance of scoring higher than a one. The effect of this cannot be underestimated in regard to assessing validity and reliability between prototype v1 and manual scoring and is an effect that was considered for future work in this area.

It should be acknowledged that the validity and reliability results outlined in the previous paragraphs may well have been affected by researcher bias during the validation testing. As previously outlined, the researcher was not blinded to the results of prototype v1, as he was required to operate the software and manually score performance simultaneously. In addition, as the researcher had developed the custom gestures for each of the three scores, prototype v1 was designed based on his interpretation of the deep squat scoring system, and as such, the chances of agreement would have been increased. Future studies where the researcher was blinded to the prototype scores, and where the scores were compared to those provided by additional certified FMS testers, provided a much greater understanding of how valid and reliable this type of software could be in automatically assessing the deep squat subtest and the remaining six FMS subtests.

4.6 DEVELOPMENT OF PROTOTYPE V2

The previous study highlighted the potential of prototype v1 to become a valid tool to automatically assess the FMS deep squat subtest. Prototype v1 could identify the 20 joints of each participant using the Kinect's Skeletal Tracking System, and in the majority of cases (133 out of 140) successfully and accurately track the movement of these joints during deep squat performance, using the Kinect's Joint Orientation System. Therefore, development work to create prototype v2 was worthwhile along with further validation work to gain a greater understanding of its validity. In particular, assessing the test-retest reliability of prototype v2 was of key importance to understand if it could be used repeatedly in a consistent and reliable manner. In addition, during the development phase of prototype v1, Microsoft released the Kinect sensor v2 which offered greater technological capabilities. These enhancements offered the possibility of improving the accuracy and reliability of prototype v2 and as such the Kinect sensor v2 was used as the hardware to support the development of prototype v2.

4.7 RESEARCH AIMS (PROTOTYPE V2)

This study attempted to answer the following research questions:

- Could software, integrated with the Kinect sensor v2, be developed that could validly assess performance on the deep squat FMS subtest compared to blinded video scoring by a certified FMS tester?

- Could software, integrated with the Kinect sensor v2, be developed that increased the test-retest reliability of the deep squat FMS subtest compared to manual screening by a certified FMS tester?

4.8 METHODS (PROTOTYPE V2 DEVELOPMENT)

4.8.1 TECHNOLOGICAL CAPABILITIES OF THE KINECT V2

A fundamental difference between the Kinect v1 and v2 was the mechanism used to measure depth information. The Kinect v1 used a method of reflecting infrared patterns against an object and back to the sensor to create a 3D depth map. The Kinect v2 used time of flight measurements, whereby it measured the time it took to send a pulsed laser signal to the object and back to the sensor to gather depth estimations. In addition, the Kinect sensor v2 offered a wider field of view than the Kinect sensor v1, with a depth range of 0.5m to 4.5m, and a 70° horizontal and 60° vertical field view, which offered greater flexibility regarding setting up FMS subtests and accurately assessing performance on same. Finally, the Kinect sensor v2 provided a greater resolution than the Kinect sensor v1, which provided greater accuracy regarding joint identification and tracking during subtest performance. Besides the colour (1920x1080@30Hz) and infrared (512x424@30Hz) data streams, it provided depth images (512x424@30Hz), body index images (512x424@30Hz) and the skeleton information for every tracked individual (25 joints@30Hz) (Muller et al. 2017). These data streams could be accessed using Microsoft's updated software development kit (v2.0) which provided additional features and functionality to support the code and algorithms written for prototype v2.

4.8.2 PROTOTYPE V2 DEVELOPMENT PHASE

Once the Kinect sensor v2 had been selected as the hardware to support the development of prototype v2, a schedule of works was devised, breaking down each area of development into distinct actions and timeframes (Table 4.5).

Table 4.5. Prototype v2 technical development plan

Task	Effort
Kinect v2 Research & Set up	3 Weeks
Raw Data Capture	2 Weeks
Machine Learning Algorithm	3 Weeks
Custom Gesture Development	2 Weeks
Initial Prototype Testing	3 Weeks
Add Initial Test Scoring	2 Weeks
Testing & Debugging	4 weeks
Total	**17 Weeks**

Kinect v2 research and set up

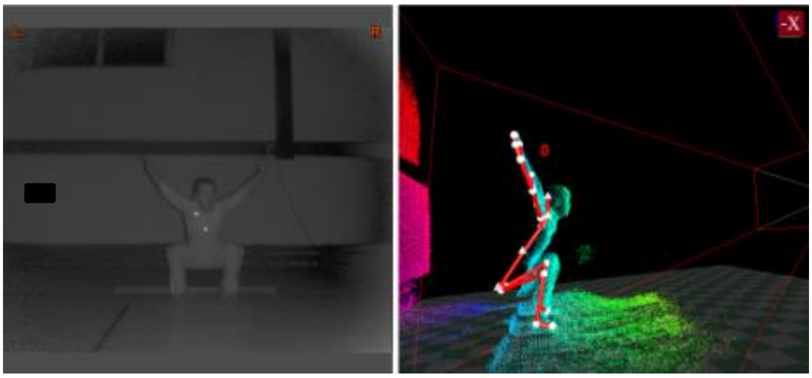
As the technical capabilities of the Kinect v2 were greater than that of the Kinect v1, research was undertaken to fully understand the key differences in functionality. In addition, time was spent working with the updated Kinect SDK to appreciate how the code written for the prototype v1 could be integrated and updated to meet the requirements of prototype v2. It was also important to understand the most effective distance and height at which to position the Kinect when recording movement of an individual, to maximise its accuracy in identifying joints and tracking joint movement. From this initial research, the most effective distance to position the Kinect away from the participant was 4.00 m in the frontal plane and at a height of 1.00 m.

Raw data capture

One of the new features identified as a positive enhancement was the Kinect sensor v2's machine learning capability. This feature could be utilised to teach the Kinect sensor v2 a deep squat score of one, two or three through the development of an algorithm that was based on raw data captured by the researcher. Raw data from 30 male participants (mean age=17.3±1.2) from the Waterford Minor hurling panel, who were medically fit to play their chosen sport at the time of testing, was collected using the Kinect sensor v2. Each participant completed all seven FMS subtests and their performances were captured as raw data. The Kinect sensor was set up 4.00 m away from the participant in the frontal plane and set on a 1.00 m high tripod. Whilst the Kinect sensor v2 was collecting the raw data, the researcher was manually recording each participant's score for each of the seven FMS subtests.

The raw data collected for the deep squat subtest generated depth and body position data (Figure 4.12 and 4.13) and was then divided into teaching and testing samples for scores of one, two and three. The teaching samples were then fed into a machine learning algorithm to learn the tagged movements for prototype v2 to be able to identify exactly what a score of one, two and three looked like based on the entire movement of the body during subtest performance.

Figure 4.12. Raw data of the deep squat set up phase as captured by the Kinect sensor v2



[redacted]

Figure 4.13. Raw data of the deep squat finish phase as captured by the Kinect sensor v2

From the initial raw data collection, the most effective use of the machine learning functionality was to assess the progress of the squat in terms of the depth that an individual could achieve with their hips through suitable knee and hip flexion e.g. how deep they were able to get their hips compared to their knees, thus assessing whether their femur was below vertical or not. The raw data analysis identified that the Kinect v2 had excellent accuracy when tracking the hip position compared to knee position in the sagittal plane and could learn the squat progress gesture without the need for additional custom gestures to be developed. This was also the most repeatable part of the deep squat movement between different individuals, with a similar pattern reproduced between participants for scores of one, two or three e.g. achieving a below parallel position for the femur signifies a score of two or three for all participants. This allowed the Kinect v2 to learn the movement which was then tested against the rest of the data to check for accuracy. As required, more teaching samples were added to the learning data and/or the tags were adjusted. This was repeated iteratively till the development team were satisfied with the learning algorithm outcomes.

Due to the large amount of variability in the position of various joints between participants when achieving the same score, it was not possible to use the machine learning algorithm to learn other gestures that make up the deep squat performance. For example, the distance between two individual’s knees could differ despite them both achieving a knee position that scores a three. This could be due to differences in body shape or size and made it difficult for the Kinect sensor v2 to accurately learn these movements using its machine learning feature.

Custom gestures

Building on the raw data captured, a custom gesture data model was created to identify individual movement errors compared to joint angle and alignment calculations that were set in accordance with the FMS scoring Manual (Functional Movement Systems, 2017). The four scoring criteria outlined in the FMS Manual formed the basis of the custom gestures developed for prototype v2 to allow the software to automatically score the deep squat subtest (Table 4.6).

Table 4.6. Custom gestures used to identify movement errors during deep squat subtest performance

Phase	Fault Code	Description	Joints			Measurement	Min	Max
PRE-TEST								
Bar on head upper arm		Checks the angle of each shoulder and that bar on head is between the min and max values	Elbow Right/Left	Shoulder Right/Left	Shoulder Left/Right	Degrees	165.00	175.00
Hand above head		Measures distance between hand and head on the Y-axis. Ensures hand is above head	Hand Right/Left	Head		Metres	0.00	0.10
START								
Hand to head distance		Measures the distance between hand and head on the X-axis on both sides and compares result	Hand Right/Left	Head		Metres	0.00	0.10
Hand behind head		Measures the distance between the hand and head on the Z-axis. Ensures hand stays behind the head	Hand Right/Left	Head		Metres	0.00	0.10
PROGRESS								
Squat progress	DS01	Using machine learnt code, determines the progress of the user in the squat				Percentage	N/A	100
Hand behind head	DS02	Measures the distance between the hand and head on the Z-axis. Ensures hand stays behind the head	Hand Right/Left	Head		Metres	0.00	0.10
Foot inline ankle	DS03	Calculates the distance on the X-axis between the ankle and foot	Foot Right/Left	Ankle Right/Left		Metres	0.00	0.03
Parallel trunk tibia	DS05	Compares the angle of the spine in the YZ plane and the angle of the tibia in the YZ plane. Min is set to 0.00. If the user is leaning back, they will generate the fault	Spine Base	Spine Shoulder		Degrees	0.00	0.01
			Ankle Right	Knee Right				
Knee over ankle	DS06	Calculates the distance on the X-axis of the knee to the ankle	Knee Right/Left	Ankle Right/left		Metres	0.00	0.03
Hip angle		This measures the angle of the joint between the two hip Joints in the XY Plane	Hip Left	Spine Base	Hip Right	Degrees	177.00	180.00
FINISH								
Hip align ankle	DS04	Measures the distance between hip joint and ankle joint in the YZ Plane	Hip Left/Right	Ankle Left/Right		Metres	0.00	0.05

For the purposes of prototype v2 assessment, the deep squat performance was broken down into four distinct phases. These phases were used to develop the relevant custom gestures (see Table 4.6) to allow prototype v2 to identify if the individual had specific joints in the correct position for each phase of performance.

1. *Pre-Test phase*

To ensure each participant assumed the correct start position, pre-test custom gestures were developed. These focussed on the position of the hands on the bar before they were fully extended overhead and gave an acceptable range of 165° to 175° on the y-axis between the elbows and shoulders. A second gesture checked that the hands were above the top of the head and gave a maximum distance of 0.10 m from the hands to the head on the y-axis.

2. *Start phase*

The start phase concentrated on the position of the arms and hands once the dowel had been fully extended overhead to ensure they remained in the correct start position. These gestures allowed a range of between 0.00 m and 0.10 m between the hands and the top of the head on the x-axis and checked that the hands were not in front of the head on the z-axis.

3. *Progress phase*

This phase incorporated six custom gestures that assessed the movement of a range of joints during the deep squat performance. This included the squat progress using the machine learning algorithm, checking if the knees and feet remained in-line with the ankles, the hands remained behind the head, the trunk and tibia remained parallel and that there was no lateral sway in either hip.

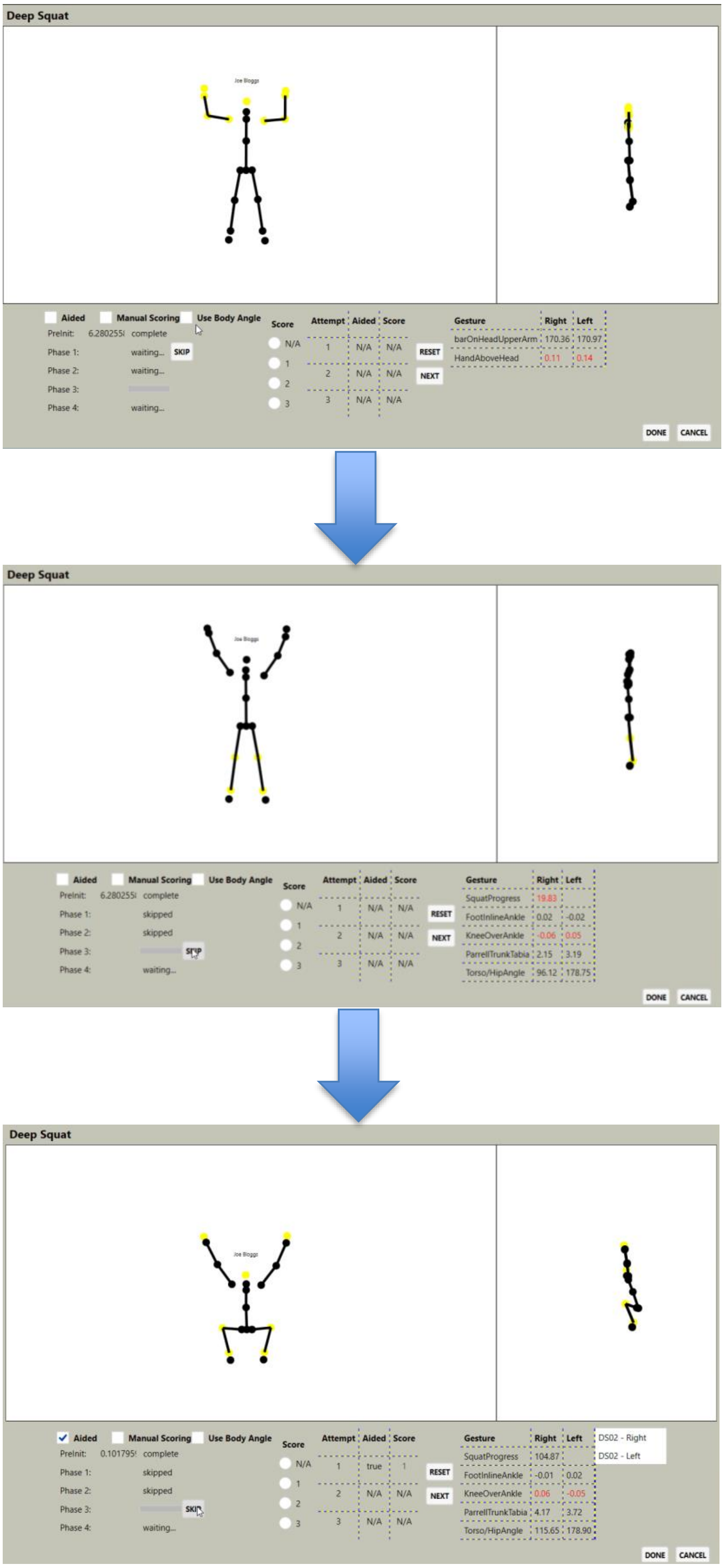
4. *Finish phase*

This phase checked for one custom gesture which considered the alignment between the hip and ankle joints at the end of the squat to ensure this had not moved beyond the 0.05 m range allowed.

The ranges given to each custom gesture were based on the researcher’s knowledge of the deep squat subtest and on internal pilot testing, whereby a range of deep squat performances were assessed by prototype v2 to check for the level of sensitivity required for each custom gesture to provide the highest level of accuracy and consistency. These gestures were then written in C-Sharp language and integrated with the Kinect SDK for the Kinect sensor v2 to be able to recognise each custom gesture and automatically score performance.

Software operation

As outlined in previous paragraphs, prototype v2 was designed to capitalise on the increased functionality offered by the Kinect v2 and build upon the work completed during the development of prototype v1. The overall aim of this development work was to attempt to increase the accuracy and reliability of prototype v2 when assessing the deep squat FMS subtest. This new functionality led to an improved user interface being developed for prototype v2 as shown in Figure 4.14.



☐ Aided ☐ Manual Scoring ☐ Use Body Angle

Prelnit: 6.280255(complete
Phase 1: waiting...
Phase 2: waiting...
Phase 3:
Phase 4: waiting...

Score: ○ N/A ○ 1 ○ 2 ○ 3

Attempt	Aided	Score
1	N/A	N/A
2	N/A	N/A
3	N/A	N/A

Gesture	Right	Left
barOnHeadUpperArm	170.36	170.97
HandAboveHead	0.11	0.14

☐ Aided ☐ Manual Scoring ☐ Use Body Angle

Prelnit: 6.280255i complete
Phase 1: skipped
Phase 2: skipped
Phase 3:
Phase 4: waiting...

Score: ○ N/A ○ 1 ○ 2 ○ 3

Attempt	Aided	Score
1	N/A	N/A
2	N/A	N/A
3	N/A	N/A

Gesture	Right	Left
SquatProgress	19.83	
FootInlineAnkle	0.02	-0.02
KneeOverAnkle	-0.06	0.05
ParrellTrunkTabia	2.15	3.19
Torso/HipAngle	96.12	178.75

☑ Aided ☐ Manual Scoring ☐ Use Body Angle

Prelnit: 0.101795(complete
Phase 1: skipped
Phase 2: skipped
Phase 3:
Phase 4: waiting...

Score: ○ N/A ○ 1 ○ 2 ○ 3

Attempt	Aided	Score
1	true	1
2	N/A	N/A
3	N/A	N/A

Gesture	Right	Left
SquatProgress	104.87	
FootInlineAnkle	-0.01	0.02
KneeOverAnkle	0.06	-0.05
ParrellTrunkTabia	4.17	3.72
Torso/HipAngle	115.65	178.90

DS02 - Right
DS02 - Left

Figure 4.14. Prototype v2 screenshots highlighting the process for recording deep squat performance

The interface included a pre-test and start set up phase, whereby the custom gestures developed were used to determine if a participant was in the correct position before the deep squat subtest began. Instead of the real-time video footage of the participant used in prototype v1, a 3D representation of the participant was used in the form of a graphical stickman highlighting the 20 joints of the body. For ease of use, any custom gesture measurements that were outside of the acceptable ranges specified were shown in red under the “Custom Gestures” table on screen and each joint incorporated within that custom gesture, highlighted in yellow on the stickman. This was designed to easily alert the operator when the pre-test positioning was incorrect.

Prototype v2 was designed to analyse the complete deep squat subtest performance as opposed to just the start and finish positions as identified by the prototype v1. This allowed prototype v2 to provide a more robust and accurate measure of performance and for the operator of prototype v2 to have greater control of the screening procedure. The operator was given the ability to begin each phase of the test themselves by clicking an onscreen button next to each subtest phase. This was designed to ensure that the software did not miss parts of the movement to keep accuracy of the software scoring at the highest possible level. As per prototype v1, the operator had the option to specify if the participant was stood on the FMS board. Once the box was checked the software was programmed to give a maximum score of two for subtest performance.

A further improvement to prototype v2 user interface, was the ability of the software to record all three subtest performances and to store each score for the operator to review at the end of the screening process. Prototype v1 was only able to record and store one subtest performance at a time, resulting in the operator having to manually record the scores given before moving on to subtest performances two and three. Prototype v2 overcame this issue, removing the need for manual recording of the scores and improving the automated nature of the scoring system.

4.9 METHODS (PROTOTYPE V2 VALIDATION)

4.9.1 SAMPLE SIZE

Development of prototype v2 took 18 weeks in total. Upon completion of the development phase, a validation protocol was undertaken to assess the validity of prototype v2 and its test-retest reliability. A sample of nine healthy male and seven healthy female participants (mean age=29.8years; SD±8.1) took part in the study. These participants were classed as physically active individuals who were currently uninjured and fit to participate in their chosen sporting activity. Ethical approval for the study was obtained from Waterford Institute of Technology Ethics committee, and each participant was asked to complete an approved Informed Consent form (Appendix A) and Pre-Test Questionnaire (Appendix B). Any participant who was unable to participate in their chosen sport at the time of screening due to injury or illness was excluded from participating in the study.

4.9.2 TESTING

Each participant was instructed to wear a T-shirt, shorts and footwear that they would normally wear to partake in exercise. The prototype v2 was set up on a Lenovo laptop running the Windows 10 operating system, which was linked to the Kinect sensor v2. The sensor was placed on a camera tripod and set at a height of 1.00 m. The FMS kit, which consisted of the FMS board and a 4ft dowel for the purposes of deep squat assessment, was set up 4.00 m from the Kinect sensor v2 in the frontal plane. Two Sony Ipela HD 1080 60p video cameras (Sony Corporation, 2013) were set up in the frontal and sagittal plane at a height of 1.00 m and distance of 4.00 m from the participant.

A research assistant controlled prototype v2 during screening, which automatically scored each participants performance on the deep squat subtest. A certified FMS tester instructed the participant in how to perform the deep squat subtest before the screening session began. As the major limitation of prototype v1 validation testing was that the researcher was not blinded to the results of the prototype scoring, it was decided that the research assistant would lead the participant through the three subtest performances, using feedback from prototype v2 to guide them. A certified FMS tester remained in attendance during subtest performance but was blinded to prototype v2 scoring. The tester observed performance to ensure each deep squat was set up and conducted as per the FMS manual, but had no input regarding scoring or the introduction of the FMS board. The video footage from the two Sony cameras were used to record deep squat subtest performance (Figure 4.15) for each participant for post testing analysis by the certified FMS tester. The researcher acknowledges that a limitation of this post video analysis technique, is that the FMS tester could see when the FMS board was used for the 2nd and 3rd attempts as instructed by the research assistant during live testing. This would have alerted them to the fact that a maximum score of two could have been given by prototype v2 for those performances. Despite this limitation, the FMS testers role was to score each performance as accurately as possible to the best of their ability, rather than attempting to match the software's score, and the protocol employed gave them every opportunity to provide all four scoring options for each participant.

Figure 4.15. Video footage captured in the frontal and sagittal plane during deep squat subtest performance

Each participant performed the deep squat three times, and each performance was automatically scored by prototype v2. The research assistant followed the onscreen instructions to correctly set the participant up prior to starting the test, and then to start and finish each test performance using the onscreen buttons. If performance was rated as less than a three by prototype v2, the research assistant instructed the participant to stand with their heels on the FMS board, and checked the "Aided" box on screen to ensure a maximum score of two could be given by prototype v2. Seventy-two hours after the first validation testing session, each participant repeated the same protocol following the same process.

The research assistant recorded the best score identified by prototype v2 from each of the three deep squat performances for each participant during each validation testing session. They recorded these scores in an excel spreadsheet that incorporated a table for each participant as per Table 4.7. The certified FMS tester reviewed the video footage taken from each testing session in the frontal and sagittal plane and provided a score for each participant based on this footage. They reviewed the video footage once per participant and in real time only to replicate live manual scoring as closely as possible. They recorded their results on a separate excel spreadsheet, and the research assistant added these scores to prototype v2 scores as per the final row shown in Table 4.7.

In addition to prototype v2 and manual scores being recorded on Table 4.7, it was also used to track the number of faults identified by prototype v2 during each testing session. Whilst prototype v2 was recording deep squat subtest performance, "Movavi" Screen Recording software was utilised to record the process. Following the inputting of all scores by the research assistant into Table 4.7, the researcher reviewed footage of each test as recorded by the "Movavi" Screen Recording software to identify exactly what faults were detected by prototype v2 during each subtest performance. Using a simple "Y" for Yes and "N" for No, the researcher noted whether each custom gesture on the user interface turned red at any point during performance to signify that the angle or distance measured had moved outside the thresholds set for the custom gesture. Despite the possible within subject variability in performance across the two testing sessions, measuring the faults identified by prototype v2 provided a good indication of test-retest reliability as well as helping to support future development work.

Table 4.7. Example of scoresheet used for validation of prototype v2

Custom Gestures	**Test 1**	**Test 2**
	Fault	**Fault**
Squat Depth (% against Benchmark)	N	N
Bar remains behind head (Yes/No)	N	N
Right foot in line with ankle (m)	N	N
Left foot in line with ankle (m)	N	N
Right knee over ankle (m)	Y	Y
Left knee over ankle (m)	Y	Y
Trunk and right tibia parallel (degrees)	N	N
Trunk and left tibia parallel (degrees)	N	N
Hip alignment (degrees)	N	N
Prototype v2 Score	**1**	**1**
Manual FMS tester score	**1**	**1**

4.9.3 STATISTICAL ANALYSIS

To ascertain if prototype v2 could accurately assess the FMS deep squat, statistical analysis was performed using IBM SPSS software version 24. Mean deep squat scores across all participants were calculated for prototype v2 and manual video scoring methods for both testing sessions. To assess the validity of prototype v2 as a method to assess deep squat performance, the absolute and percentage agreement between the prototype and manual video scoring methods for all participant's scores were calculated. Additionally, a Spearman rank-order correlation test was conducted to understand the level of correlation between prototype v2 and manual video scoring, as another method to test the validity of the software. To analyse test-retest reliability of prototype v2 an Intraclass correlation coefficient (ICC) statistic was calculated for both the deep squat subtest score and the number of errors detected by prototype v2 for each testing session using an alpha level of 0.05. ICC values of 0.75 and above represent good reliability, those between 0.50 and 0.74 represent moderate reliability, and those below 0.50 indicate poor reliability (Teyhan et al., 2012).

4.10 RESULTS (PROTOTYPE V2)

Sixteen participants performed the deep squat subtest three times during two separate testing sessions, 72 hours apart. For testing session one, mean deep squat subtest score was 1.3 (SD±0.6) for manual video scoring and 1.2 (SD±0.4) for prototype v2 scoring. For testing session two, mean deep squat subtest score was 1.4 (SD±0.6) for manual video scoring and 1.2 (SD±0.4) for prototype v2 scoring. When considering both testing sessions, absolute agreement between manual video and prototype v2 scoring was good for scores of one and two, but there was no agreement for scores of three (Table 4.8).

Table 4.8. Breakdown of absolute agreement between prototype v2 and manual video scoring for scores of one, two and three on the deep squat FMS subtest. Figures in red highlight disagreement between methods

	Manual Scores			
Software Scores	**1**	**2**	**3**	**Total**
1	22	2	2	26
2	0	6	0	6
3	0	0	0	0
Total	22	8	2	32

Overall agreement for all deep squat performances across both testing sessions was 87.5%. Scores of one (n=26) reported 86.6% agreement and scores of two (n=8) reported 75% agreement between the scoring methods. All 22 deep squat performances identified by manual video scoring as a one, were identified automatically by prototype v2 as the same score, whilst six of eight deep squat performances identified by manual video scoring as a two were also identified automatically by prototype v2 as the same score (Figure 4.16). Two deep squat performances scored as a three by manual video scoring were incorrectly given a score of one by the automated prototype v2 scoring system. When considering concurrent validity for testing session one, there was a moderate positive correlation between manual video scoring and prototype v2 scoring (r_s=0.51; p=0.04). For testing session two, the level of correlation between the two scoring methods reduced to (r_s=0.37; p=0.16).

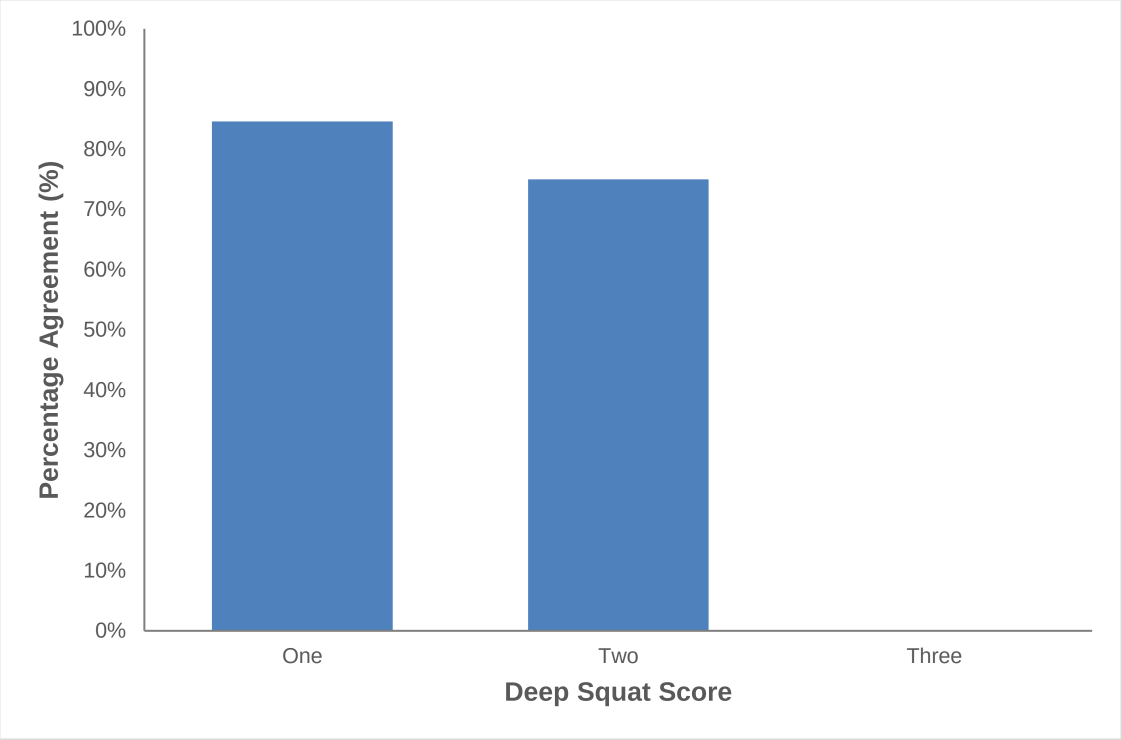


Figure 4.16. Summary of percentage agreements between prototype v2 and manual video scoring methods for the deep squat FMS subtest (n=32)

From a test-retest reliability perspective, prototype v2 showed moderate reliability between testing sessions, with an ICC of 0.74 (95% CI, 0.26 to 0.91; p<0.01). This test-retest reliability is lower than that seen for manual video scoring between the same two testing sessions with a reported ICC of 0.91 (95% CI, 0.76 to 0.97; p<0.01) (Figure 4.17).

When considering the number of faults identified by prototype v2 between the two different testing sessions, test-retest reliability was good with an ICC of 0.79 (95% CI, 0.40 to 0.93; p<0.01).

Figure 4.17. Test-retest reliability of scoring methods used to assess deep squat performance

4.11 DISCUSSION (PROTOTYPE V2)

The results indicated that prototype v2 correctly identified 28 out of 32 deep squat performances across two testing sessions, compared to manual video scoring, with an overall agreement of 87.5%. This overall agreement level is lower than that reported for prototype v1 (95%) and by some previous research in this area (Schneiders, Davidsson, Hörman, & Sullivan, 2011a), but remains higher than that seen in additional research that assessed agreement between manual raters (Minick et al., 2010; Teyhen et al., 2012) and analysed the validity of a semi-automated system used to score the deep squat subtest (Jensen, Weilbrenner, Rott, & Eskofier, 2013). However, there was only a moderate correlation (r_s=0.51) between automated and manual scoring for testing session one and low correlation (r_s=0.37) between the two methods for testing session two. Therefore, the validity of using prototype v2 to accurately assess deep squat FMS subtest performance could only be classified as moderate despite high levels of absolute agreement.

The level of detail incorporated into the prototype v2 when assessing deep squat performance far exceeded that seen in prototype v1. This included the tracking of 20 joints compared to only six in prototype v1 and measuring both range of motion and alignment between joints in 3D space, compared to just range of motion at six joints in the frontal and sagittal plane for prototype v1. The increased level of sensitivity for prototype v2 led to more faults being recognised during deep squat performance but did not increase the accuracy of prototype v2's scoring compared to manual assessment. Upon analysis of the results, two of the custom gestures incorporated into prototype v2 were too sensitive and led to a large majority of the disagreement between the scoring methods. The "knee align ankle" and "hip align ankle" on both the right and left sides were the two custom gestures that recorded more faults than any other. For the "knee align ankle" the range of movement was set at 0.03 m and for the "hip align ankle" was 0.05 m. Upon analysis of the video footage after all results had been collated, it was possible for the alignment between the knees and ankles, and hip and ankles, to move outside these thresholds but the participant still receive a score of two or three from manual scoring assessment. In these instances, prototype v2 would indicate a fault and score the participant accordingly resulting in a disagreement with manual video scoring. This was the case for the two participants that were manually scored as a three but were scored as a one by prototype v2 as the alignment of the knees and ankles moved outside the custom gesture range that was set for prototype v2. Future development of the software needed to carefully consider these custom gestures and either change the measurement used or increase the ranges applied to improve agreement levels with manual scoring.

Only 25% of performances resulted in a manual score of two. This is a relatively low percentage compared to previous research that has reported 50% or more of participants scoring a two (Schneiders et al., 2011; Teyhen et al., 2012). Although it is difficult to establish the cause of this, the age of the cohort in these studies could have had an impact. The mean age of the 16 participants in the current study was 29.8 years compared to 21.9 years in the study by Schneiders et al. (2011) and 25.2 years in the study by Teyhen et al. (2012). Although research to date has not looked at the effect of increased age on FMS score, the deep squat is a complex movement that involves multi-joint movement, and it would not be unexpected for older participants to perform worse on this subtest. Research on the deep squat by Butler, Plisky, Southers, Scoma, & Kiesel (2010) identified that mobility and/or stability limitations at the ankle joint could inhibit performance on the deep squat. They suggested that improvements in hip, knee and ankle mobility can improve performance on the subtest from a one to a two. As the current study used an older cohort, it could be suggested that the level of mobility in these key joints had decreased as the participants had got older, which may have contributed to the low number of twos recorded.

The test-retest reliability of prototype v2 was moderate (ICC=0.74) and was lower than the test-retest reliability of manual video scoring (ICC=0.91). However, this level of test-retest reliability is comparable to that seen in previous research assessing the reliability of manual scoring (Shultz et al., 2013; Teyhen et al., 2012). It should also be noted that manual video scoring identified within subject variability across the deep squat performances of two participants. These were also correctly identified by prototype v2, although the system did record an additional change in performance score for a third participant as well. When conducting the same testing protocol on two separate occasions only 72 hours apart, there is a possibility of some learning effect for the participants that may have led to improved performance during the second testing session. If this did occur, it would have adversely affected the test-retest reliability results. Interestingly, the level of test-retest reliability of prototype v2 increases when considering the number of faults identified (ICC=0.79). When analysing the faults in more detail there is an excellent level of consistency in the faults identified and the score provided by prototype v2 e.g. any score of two provided by prototype v2 for both testing sessions correctly attributed no faults to those performances. This suggests that if the sensitivity of the custom gesture thresholds could be adjusted to match that of manual scoring, prototype v2 would have the level of consistency required to ensure test-retest reliability to match or better manual video scoring. However, based on the current study, prototype v2 can only be considered to have moderate test-retest reliability which is not as high as that seen when using manual video scoring.

Incorporating manual video scoring into the study ensured the FMS tester remained blinded to prototype v2 results at all times which enhanced the quality of this study compared to that conducted for prototype v1. Yet, incorporating video scoring reduced the applied nature of the study as the likelihood of a practitioner using video scoring in the field is lower than using live testing due to the practicalities involved. One of the main aims of developing the new automated technology to assess the FMS was to increase the speed of the screening process and make the whole protocol as simple as possible for the practitioner. As such it could be argued that comparing prototype v2 results to live manual scoring would provide a more realistic measure of test-retest reliability and accuracy from an applied perspective. Minick et al. (2010) reported that the level of inter-rater reliability between novice raters was higher than that seen for expert raters when using video scoring to assess FMS performance. Parenteau et al. (2014) also reported excellent intra-rater reliability between video scoring on two separate occasions by two expert raters. These studies indicate that video scoring offers a more precise and reliable method by which to score the FMS, but does not replicate the real world situation that most practitioners find themselves in when assessing the FMS. Smith, Chimera, Wright, & Warren (2013) make the point that using video scoring does not replicate real time administration, whereby the rater must observe the performance from multiple viewpoints to ascertain the correct score. With video scoring only offering frontal and sagittal plane views, the likelihood of a fault being missed is increased. Future development work in this area should consider comparisons against both live and video manual scoring to fully understand the level of validity and reliability that the automated software offers. Consideration should also be given to a comparison against alternative, objective, measurement devices that remove the subjectivity of manual scoring.

4.12 CONCLUSION

These research studies have outlined the development and validation processes that were undertaken to design prototype v1 and prototype v2 software that utilised the Kinect sensor v1 and v2 to automatically assess the FMS deep squat subtest. Results from study one suggested that the prototype v1 had the potential to offer a solution to reduce the current subjectivity seen when scoring the deep squat subtest. It was difficult to accurately assess the true level of validity and reliability of prototype v1 due to the limitations of the study and relatively basic nature of the software developed. However, the study showed that further research to build on the work completed to develop prototype v1 was justified, as the technology had the ability to track the key joints of the body and, in most cases, could track the movement of these joints during deep squat performance.

The focus of study two was on the development of additional custom gestures to assess the range of motion at more joints and for the alignment of key joints to be measured. Utilisation of the additional capabilities provided by the Kinect sensor v2, allowed for advanced joint identification and tracking to be developed within prototype v2, to provide a more thorough assessment of deep squat performance. Furthermore, study two incorporated blinded manual assessment of deep squat performance and assessed the test-retest reliability of prototype v2's scoring system at different time points.

The validity and test-retest reliability of prototype v2 was moderate compared to manual video scoring, with good overall absolute agreement between the two scoring methods. Prototype v2 showed a high level of test-retest reliability regarding the number of faults identified across the two testing sessions and in attributing the correct score against the faults highlighted. This suggests that the algorithms developed using the FMS Manual (Functional Movement Systems, 2017) scoring criteria to teach prototype v2 the difference between scores of one, two and three, worked effectively. It was found that certain custom gestures were over sensitive when considering the thresholds applied, which led to prototype v2 incorrectly identifying faults in certain performances. By adjusting these thresholds or changing the custom gestures, it would be expected that the level of validity and reliability of prototype v2 would increase.

The results from the two studies suggest that future work to develop the prototype's ability to automatically assess additional FMS subtests was warranted. The level of validity and reliability identified to date is comparable to that seen in manual scoring of the deep squat FMS subtest and as such it should be possible to replicate and improve on this for the remaining six FMS subtests. Further work considered the most effective set up position for the Kinect sensor v2 when assessing FMS subtests. Study two successfully showed that it was possible to increase the range of prototype v2 to 4.00 m and height to 1.00 m. Given that some of the additional FMS subtests require the participant to be in a prone or supine position, further adjustments to the height were required to enhance accuracy. It was also necessary to move the position of the sensor from the frontal position for subtests that involved more complex multi-joint movements. Future studies manipulated these set up parameters to understand how these could positively affect validity and reliability.

Finally, further validation of the prototype considered the use of both live and video manual scoring against which to compare the prototypes results. Although the video scoring method appears to be the most reliable in terms of correctly scoring FMS subtests, it does not fully replicate applied settings where live scoring is more common. Therefore, for the prototype to be considered a valid and reliable tool to use to assess the FMS, a comparison against both methods was made to provide a full evaluation of its capabilities.

CHAPTER FIVE

DEVELOPMENT OF MOTION TRACKING SOFTWARE TO ASSESS THREE FMS SUBTESTS

5.0 INTRODUCTION

Chapter 4 outlined the development and validation work that was completed using the Kinect v1 and v2 depth cameras as the hardware to support the novel software developed to create prototype v1 and v2. The development of prototype v1 and v2 focussed entirely on the deep squat FMS subtest, with the objective of developing software that could automatically score the deep squat subtest in a valid and reliable manner. The previous chapter showed that prototype v2 had good levels of agreement when compared to manual scoring, and also reported similar levels of validity and reliability compared to previous research that has compared manual raters using the FMS (Shultz et al., 2013; Teyhen et al., 2012). Hence, prototype v2 showed promise as an automated tool to assess the deep squat subtest and further work was warranted to improve its levels of reliability and validity, as well as its ability to assess the remaining six FMS subtests.

5.0.1 CUSTOM GESTURES

The next phase of development work focussed on three key areas to attempt to progress the software to reach a level where it could validly and reliably be used in applied settings. Firstly, the custom gestures developed for each subtest were carefully considered to ensure the specific requirements of each subtest were met. The results from prototype v1 and v2 suggested that the custom gestures developed for the deep squat were overly sensitive and as such led to participants being underscored compared to manual rating. Whilst the custom gestures were designed to replicate the scoring criteria provided in the FMS Manual, they were also able to track much finer joint movements than the naked eye, which gave greater opportunity for the prototype to find errors in performance. For the remaining six FMS subtests, the scoring criteria for each was considered, so the movement could be broken down into suitable phases for the software to be able to accurately score performance. The role machine learning could play in assessing each subtest was also considered, as the deep squat subtest highlighted that this functionality was not suitable to use for every movement.

5.0.2 KINECT SENSOR SET UP

The set-up of the Kinect sensor needed to change for different subtests to maximise the opportunity for the sensor to identify joint centres and track them during performance of a certain subtest. The previous set up of 1.00 m high and 4.00 m distance from the participant in the frontal plane worked well for the deep squat subtest. However, for those subtests that involve the participant being in a prone, supine or quadruped position (active straight leg raise, trunk stability and rotary stability subtests) the height of the sensor needed to be adjusted to increase the ability of the sensor to identify joint centres even when the participant was not facing the sensor. Also, for the subtests that required an almost 360° view of performance (inline lunge and rotary stability subtests) placing the sensor in the frontal plane may not have been the most effective positioning to maximise accuracy of scoring. As such, the development work moving forward assessed the most appropriate height and position for the sensor for all seven FMS subtests.

5.0.3 SCORING METHODS

Finally, both blinded live and video manual scoring methods were incorporated into future studies to compare against automatic software scores to assess validity and reliability. As the software was designed to be used in applied settings, ensuring that it could score the FMS more reliably than live manual scoring was important. However, it was also accepted that video scoring is a reliable method to use to score the FMS (Minick et al., 2010; Shultz et al., 2013), yet is obviously less practical and more time consuming to complete and therefore less commonly used in applied settings. By assessing the reliability and validity of the software against both live and video scoring, it was possible to establish if the software was comparable to both methods and therefore understand if it was a valid method to use no matter what environment the FMS was conducted in.

Following on from the work completed on prototype v1 and v2, this programme of research was split into two distinct studies. Study three, discussed in this chapter, focussed on three of the FMS subtests, the deep squat, hurdle step and inline lunge. The aim of the study was to assess if the novel software, linked with the Kinect v2 sensor, could reliably and validly assess performance on the three subtests compared to manual scoring. This allowed assessment of how easily the code and algorithms already developed for the deep squat could be applied to the hurdle step and inline lunge subtests, and if it was possible to overcome any major obstacles that appeared. The two extra subtests were chosen as they were the next two subtests in the FMS process and because they did not require a participant to be led in a prone or supine position on the floor. As such they provided similar challenges to the deep squat subtest and helped provide additional information that could be used to improve future development work on the remaining FMS subtests.

5.1 RESEARCH AIMS (PROTOTYPE V3)

This study attempted to answer the following research question; Could software, integrated with the Kinect sensor v2, be developed that could validly and reliably assess performance on the deep squat, hurdle step and inline lunge FMS subtests compared to manual scoring by a certified FMS tester?

5.2 METHODS (PROTOTYPE V3 DEVELOPMENT)

5.2.1 PROTOTYPE V3 DEVELOPMENT PHASE

Following completion of the development and validation phases for prototype v2, the results and errors that had been identified within the custom gesture framework were reviewed to better understand how improvements to prototype v3 could be made. A development plan was devised that was designed to integrate the two new subtests, as well as a database to allow the software to store team, session and participant information, into the existing user interface (Table 5.1).

Table 5.1. Prototype v3 development plan

Task	Effort
Raw Data Review for HS/IL	2 Weeks
Machine Learning Algorithm for HS/IL	2 Weeks
Custom Gesture Development	3 Weeks
Database development	2 Weeks
Initial Prototype Testing	2 Weeks
Add Initial Test Scoring	2 Weeks
Testing & Debugging	3 weeks
Total	**16 Weeks**

Raw data review

Raw data captured during the development of prototype v2, outlined in section 4.8.2, was used to assess whether a machine learning algorithm could be created for both the hurdle step and inline lunge subtests. As per the deep squat, the footage was reviewed to select scores of three, two and one to use as teaching samples to teach the Kinect sensor the movements that related to each of the scores available.

Machine learning algorithm

Once these teaching samples were selected and fed into the machine learning algorithm it was found that the level of variation across performances that achieved the same score was too wide for the Kinect to be able to correctly identify the movements as a score of three, two or one. This was found for both the hurdle step and inline lunge subtests. When using this method for the deep squat, the depth of the squat could be tracked by the Kinect sensor as it was clearly able to identify the hip and knee joints, to determine if the hip had moved below the height of the knee. It could therefore be taught that a performance which moved the hip below the knee was a three (for that custom gesture) or a two if the “aided” button had been checked. This was a relatively simple process as it involved only two joints and the parameters of performance were the same for every individual. For the hurdle step and inline lunge, it was not possible to identify a specific part of the movement that could utilise the machine learning function in the same way. Both subtests involved multiple joints and assessment of both upper and lower limbs and have a progression of movement which is similar to the depth of the deep squat. However, the way to measure this progression involved a body part touching an inanimate object, which the Kinect is not able to accurately identify. For the hurdle step the heel of the moving foot must touch the ground once it has reached over the rubber band. For the inline lunge, the back knee must touch the FMS 6ft by 2ft board, before returning to the start position. If the Kinect could accurately identify these inanimate objects, these gestures could have been used to develop a machine learning algorithm for each subtest. However, it was not possible for the Kinect to reliably identify inanimate objects and as such machine learning was not used for the hurdle step or inline lunge subtests.

Custom gesture development

For each of the three subtests, custom gestures were developed to allow prototype v3 to assess the range of motion at key joints and the alignment between these joints to score overall performance. For the deep squat certain gestures were adjusted and others removed that provided no benefit to the scoring of the deep squat subtest, based on feedback received during prototype v2 validation. For the hurdle step and inline lunge subtests, the scoring criteria outlined in the FMS Manual were used and broken down into specific ranges of movement to distinguish between scores of one, two and three.

Deep Squat

Following completion of the validation testing for prototype v2, video footage of all 32 deep squat performances were reviewed to assess the custom gesture measurements identified by prototype v2. This was designed to establish the level of sensitivity of the custom gesture measurement ranges (angle or distance) to understand if any incorrectly identified a fault and hence provided an incorrect score. The review also established whether a custom gesture was adding value to performance assessment or in fact was causing errors to be identified that were not actually present.

When considering the sensitivity of each of the custom gestures it was clear that three of the custom gestures had thresholds that were overly sensitive and causing prototype v2 to underscore performance compared to manual scoring. For the “knee over ankle” and “foot inline ankle” gestures, the analysis showed that the angle ranges set were too sensitive. As a result, for two of the deep squat performances scored as a three by manual scoring, prototype v2 scored performance as a one as the angles identified were outside the acceptable thresholds set for both custom gestures. This highlighted that both the knees and feet could move slightly during performance and still stay within the FMS scoring criteria and therefore the thresholds for these two custom gestures needed to be increased. For the “parallel trunk tibia” it was identified that this gesture was the cause of two incorrect scores being provided by prototype v2 compared to manual scoring. This custom gesture was designed to measure the alignment between the trunk and tibia in the sagittal plane to check that they remained parallel during squat performance. With the thresholds set as they were for prototype v2, any slight movement of the trunk was causing an error to be identified, when in fact the movement was not large enough to result in the manual score being adjusted. Figure 5.2 shows the custom gesture list used for prototype v3 and the adjustments that were made to specific custom gestures.

Table 5.2. Custom gestures used to identify movement errors in prototype v3 during deep squat subtest performance

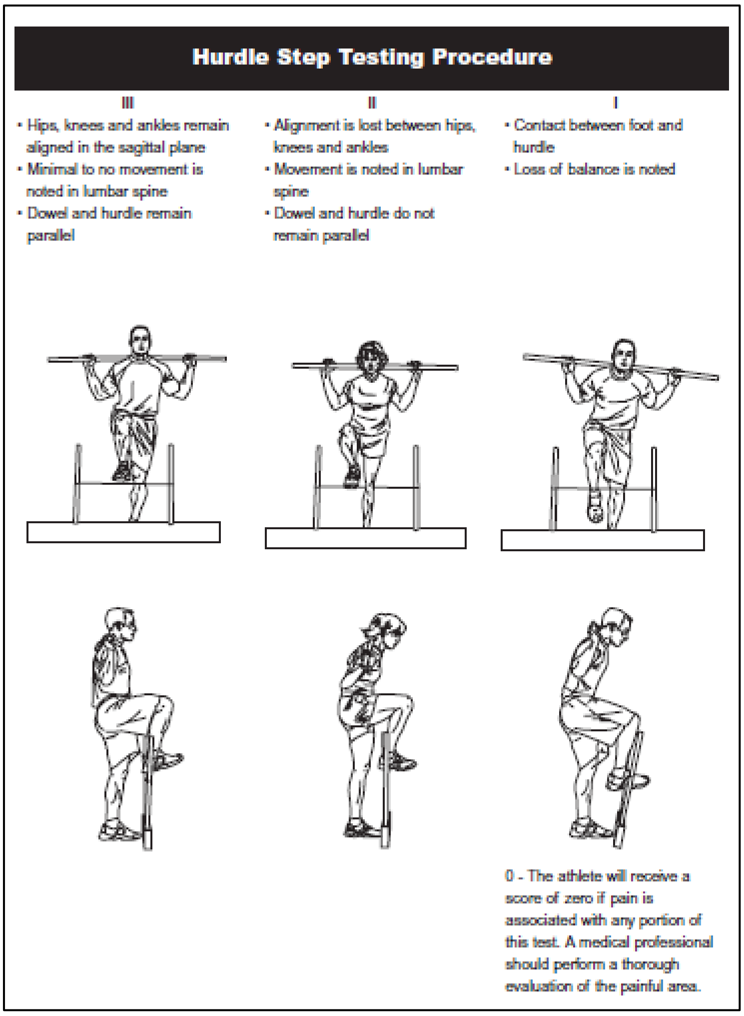
Phase	Description	Joints			Measurement	Min	Max
PRE-TEST							
Bar on head upper arm	Checks the angle of each shoulder and that bar on head is between the min and max values	Elbow Right/Left	Shoulder Right/Left	Shoulder Left/Right	Degrees	165.00	175.00
Hand above head	Measures distance between hand and head on the Y-axis. Ensures hand is above head	Hand Right/Left	Head		Metres	0.00	0.10
START							
Hand to head distance	Measures the distance between hand and head on the X-axis on both sides and compares result	Hand Right/Left	Head		Metres	0.00	0.10
Hand behind head	Measures the distance between the hand and head on the Z-axis. Ensures hand stays behind the head	Hand Right/Left	Head		Metres	0.00	0.10
PROGRESS							
Squat progress	Using machine learnt code, determines the progress of the user in the squat				Percentage	N/A	100
Hand behind head	Measures the distance between the hand and head on the Z-axis. Ensures hand stays behind the head	Hand Right/Left	Head		Metres	0.00	0.10
Foot inline ankle	Calculates the distance on the X-axis between the ankle and foot	Foot Right/Left	Ankle Right/Left		Metres	-0.02	0.03
Parallel trunk tibia	Compares the angle of the spine in the YZ plane and the angle of the tibia in the YZ plane. Min is set to 0.00. If the user is leaning back, they will generate the fault	Spine Base	Spine Shoulder		Degrees	30.00	N/A
		Ankle Right	Knee Right				
Knee over ankle	Calculates the distance on the X-axis of the knee to the ankle	Knee Right/Left	Ankle Right/left		Metres	-0.02	0.03

When considering gestures that were causing errors to be identified, two gestures were found to be constantly incorrect. The “hip angle” and “hip align ankle” gestures often identified faults when in fact no faults were present when manually scoring performance. The “hip angle” measured the alignment between the left/right hip and the spine base as identified by the Kinect “Skeletal Tracking” system, essentially measuring lateral sway of the trunk during performance. Upon review of prototype v2 performances, during many of the performances scored as a one, prototype v2 identified a fault with this gesture, which was not seen during manual scoring. Although this did not impact the score given by prototype v2, due to other gestures correctly being identified as a fault, this gesture had the potential to cause accuracy issues in the future. During further internal pilot testing of prototype v3, despite amendments to the range of this gesture, prototype v3 was still identifying errors when none were present according to manual scoring. Therefore, this gesture was removed from the scoring system altogether to avoid further problems during future validation testing.

The “hip align ankle” gesture was designed to check if the hips remained in line with the ankle before, during and after performance. This was designed to ensure that any misalignment in the hips during deep squat performance would be identified. Upon review, it was found that many individuals do not start the deep squat with their hips in line with their ankles due to their natural body shape, and as such prototype v2 would never be able to give them a score of two or three whilst this custom gesture was operational. In addition, hip misalignment during deep squat performance often leads to knee misalignment and therefore any movement error would be correctly identified by this custom gesture and the performance correctly scored. The “hip align ankle” gesture was therefore removed from the custom gesture list used for prototype v3.

Hurdle Step

The hurdle step subtest challenges each participant’s unilateral balance, co-ordination and stability. A full description of all seven FMS subtests can be found in Appendix F. The FMS Manual provides scoring criteria for the hurdle step subtest as outlined in Figure 5.1.



Hurdle Step Testing Procedure

III
- Hips, knees and ankles remain aligned in the sagittal plane
- Minimal to no movement is noted in lumbar spine
- Dowel and hurdle remain parallel

II
- Alignment is lost between hips, knees and ankles
- Movement is noted in lumbar spine
- Dowel and hurdle do not remain parallel

I
- Contact between foot and hurdle
- Loss of balance is noted

0 - The athlete will receive a score of zero if pain is associated with any portion of this test. A medical professional should perform a thorough evaluation of the painful area.

Figure 5.1. The hurdle step FMS subtest scoring criteria (Functional Movement Systems, 2017)

As Figure 5.1 highlights, there are three main scoring criteria for the hurdle step subtest:

- Alignment between hips, knees and ankles in sagittal plane
- Movement in the lumbar spine
- Dowel and hurdle remain parallel

The hurdle step subtest also has a very specific set up procedure to ensure the test is conducted in a consistent and reliable manner every time. This includes measurement of each participants tibial tuberosity bone prior to subtest performance to establish exactly what height to set the hurdle using a rubber band. Based on this set up procedure and the scoring criteria outlined above, a set of custom gestures were developed for prototype v3 that allowed hurdle step performances to be scored automatically based on the FMS Manual guidelines. Table 5.3 provides details of these custom gestures.

Table 5.3. Custom gestures used to identify movement errors in prototype v3 during hurdle step subtest performance

Hurdle Step subtest	Description	Units	Min	Max
1. Set Up Phase				
Hand to Neck Distance	Compares distance between hand and neck on each side	Metres	0.00	0.01
2. Test Phase				
Hip Extension	Measures the angle of the hip in YZ Plane	Degrees	87.00	95.00
Active Leg Knee Alignment	Measures the distance on X-axis of knee and hip on active leg	Metres	0.00	0.05
Active Leg Ankle Alignment	Checks distance on X-axis of ankle and knee on active leg	Metres	0.00	0.05
Lateral Movement	Measures the angle of the spine on X-axis	Degrees	85.00	95.00
Upper Body Rotation	Compares distances on the Z-axis of shoulder to hip on each side	Metres	0.00	0.05

To allow prototype v3 to automatically assess hurdle step performance it was necessary to break the movement down into two distinct phases which incorporated different custom gestures. These phases ensured each participant was set up correctly prior to each performance before the software began to assess performance.

1. Set Up Phase

This phase identified if the participant's hands were correctly positioned with the dowel resting on the shoulders and the hands placed an equal distance from the neck on each side. The range set for this gesture was 0.00 m to 0.01 m to provide a very sensitive measure, ensuring set up was correct prior to the start of the subtest.

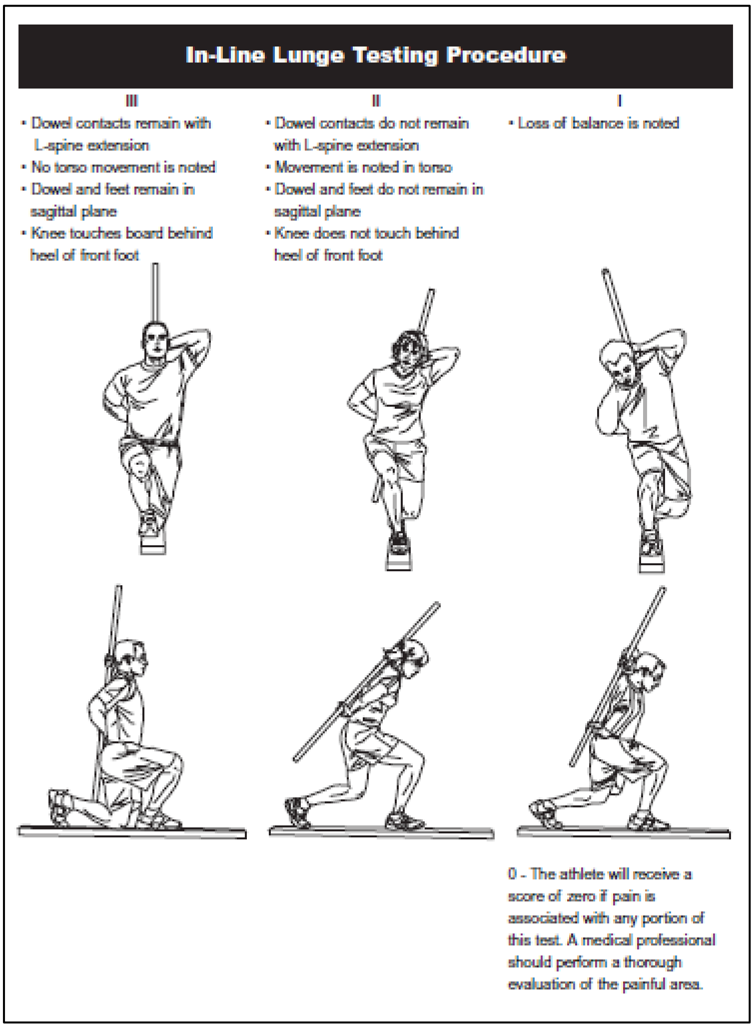
2. Test Phase

Once set up was complete, the software moved into the test phase which incorporated five custom gestures. Firstly, it assessed hip extension to determine if there was any movement of the trunk during performance. The range was set between 87° and 95° on the z-axis to allow a small amount of movement forward and backwards in the trunk. The next two custom gestures assessed the leg that was moving to ensure it moved in a controlled manner without compensation. It assessed the alignment between the hip and the knee and the ankle and the foot on the active leg using a range for both of 0.00 m to 0.05 m on the x-axis. A common compensation when the leg is active is rotation in the hip and ankle, and these gestures were designed to check for these common errors. The final two gestures assessed the movement of the trunk during performance. Lateral movement of the trunk was assessed using a range of 85° to 95° in the x-axis and upper body rotation was measured using alignment of the shoulders and hips on the z-axis with a range of 0.00 m to 0.05 m.

Manual features relating to recording the height of the tibial tuberosity bone, whether the participant felt pain and if the hurdle was hit during performance, were incorporated into the user interface. The software operation procedures for each subtest are discussed in more detail in Appendix H.

Inline Lunge

The inline lunge subtest assesses an individual's mobility, stability and balance whilst the feet are in a scissor position. The FMS manual provides clear scoring criteria for the inline lunge subtest as outlined in Figure 5.2.



In-Line Lunge Testing Procedure

III
- Dowel contacts remain with L-spine extension
- No torso movement is noted
- Dowel and feet remain in sagittal plane
- Knee touches board behind heel of front foot

II
- Dowel contacts do not remain with L-spine extension
- Movement is noted in torso
- Dowel and feet do not remain in sagittal plane
- Knee does not touch behind heel of front foot

I
- Loss of balance is noted

0 - The athlete will receive a score of zero if pain is associated with any portion of this test. A medical professional should perform a thorough evaluation of the painful area.

Figure 5.2. The inline lunge FMS subtest scoring criteria (Functional Movement Systems, 2017)

As figure 5.2 highlights, there are four main scoring criteria for the inline lunge subtest:

- The dowel remains in contact with the body
- The knee touches the board behind front heel
- No torso movement
- The dowel and front foot remain in the sagittal plane

These four criteria were used as the basis for the custom gestures that were written for prototype v3 to be able to establish scoring errors during performance as highlighted in Table 5.4.

Table 5.4. Custom gestures used to identify movement errors in prototype v3 during inline lunge subtest performance

Inline Lunge subtest	Description	Unit	Min	Max
1. Test Phase				
Body Lean	Checks torso angle	Degrees	85.00	95.00
Heel Flat	Checks front heel remains flat on the board	Metres	0.00	0.01
Knee and Ankle Inline	Checks distance of knee and ankle on the Z-axis	Metres	0.00	0.05
Lateral Movement	Measures the angle of the spine on X-axis	Degrees	85.00	95.00
Upper Body Rotation	Compares distance between shoulder and hip on the X-axis	Metres	0.00	0.02

The inline lunge subtest had a relatively complex set up procedure which involved the feet being placed the correct distance apart and the hands being placed in the correct position behind the back. Despite attempts to automatically assess this using prototype v3, it was not possible to complete this in a reliable and accurate manner because the sensor could not identify the hands behind the back or accurately assess the distance the feet were apart. As such the set-up phase was removed altogether from the custom gestures for the inline lunge. Therefore, the set up of a participant for the inline lunge was a fully manual process and prototype v3 assumed this had been completed correctly prior to the performance beginning.

As the inline lunge required assessment from 360° placing the sensor in the frontal plane may not have offered the level of accuracy required to score the test correctly. Therefore, during pilot testing, the FMS board was set at different angles to assess which allowed the Kinect sensor to identify all joint centres and accurately track movement during test performance. The board was set at four different angles from the Kinect sensor, 30°, 45°, 65°, and 90°. The ability of the Kinect to track joints whilst an individual performed the inline lunge subtest with the board at these four different angles was assessed, and it was determined that setting the board at a 30° angle from the Kinect sensor provided the most accurate position for joint tracking. The custom gestures that were developed for the inline lunge were written based on the board being set at this angle during inline lunge testing.

The test phase included five custom gestures based on the FMS manual scoring criteria aimed at maximising the accuracy of prototype v3 scoring. The first custom gesture, body lean, assessed the angle of the torso, using an acceptable range of 85° to 95° of movement forward or backwards. The next gesture assessed whether the heel of the front foot is raised during performance. A sensitive range of 0.00 m to 0.01 m was set as any slight raise of the heel is deemed as an error when scoring this performance manually. The alignment of the front knee and ankle were then assessed in the third gesture, to check on the stability of the knee. Based on the feedback from the deep squat subtest a range of 0.00 m to 0.05 m was set, as a small amount of movement of the knee would be deemed acceptable when manually scoring performance. The final two gestures used the same measures that were incorporated into the hurdle step subtest. They assessed the lateral movement of the trunk and rotation of same using the ranges of 85° to 95° and 0.00 m to 0.02 m respectively.

As per the hurdle step subtest some manual checks were included within the inline lunge user interface to identify if the participant was in pain, whether the knee has touched the board and whether the dowel has remained in contact with body.

Software Operation

Each of the three subtests were designed to be as automated as possible and to allow a user to conduct a subtest in an efficient and consistent manner. Upon analysis of prototype v2, it became apparent that the software needed to provide the user with a method to control the start of each subtest once the participant had managed to get in the correct set up position. Once this had occurred, prototype v2 automatically moved onto the testing phase and started to assess performance. Although this worked efficiently for the deep squat, it was envisaged this may cause problems for other subtests. Those subtests that required the tester to help the participant get set up manually for performance, such as the inline lunge, would need to allow the user to control when the software began assessment. Therefore, once the set-up phases were complete, a pop up window was added that required the user to click “OK” before test assessment began. This was installed for all three subtests in prototype v3. A detailed breakdown of how to operate the user interface for each subtest can be found in Appendix H.

5.3 METHODS (PROTOTYPE V3 VALIDATION)

5.3.1 SAMPLE SIZE

Upon completion of prototype v3 development, which took 16 weeks, validation testing was undertaken to assess the validity of prototype v3 compared to manual scoring of the three FMS subtests. Ethical approval for this study was granted by the Waterford Institute of Technology Ethics Committee. All participants were informed of the benefits and risks of the investigation prior to signing an institutionally approved Informed Consent form (Appendix A) to participate in the study. Twenty-seven healthy male participants (mean age:19.8±3.8 years) took part in the study. The participants were members of the Wexford Youths League of Ireland senior soccer team and the Killester senior basketball team. Prior to participation each subject completed a Pre-Test Questionnaire (Appendix B). Any participant unable to participate in their chosen sport at the time of screening due to injury or illness was excluded from participating in the study.

5.3.2 TESTING

Each participant was instructed to wear a T-shirt, shorts and footwear that they would normally wear to partake in exercise. Prototype v3 was set up on a Lenovo laptop running the Windows 10 operating system, which was linked to the Kinect sensor v2. The sensor was placed on a camera tripod and set at a height of 1.00 m. The FMS kit, which consisted of a 6ft by 2ft board, a 4ft dowel and a rubber cord, was set up 4.00 m from the Kinect sensor v2 in the frontal plane. For the inline lunge subtest, the FMS board was turned to an angle of 30° from the Kinect sensor.

A research assistant controlled prototype v3 during screening which automatically scored each participants performance on the three subtests. Prior to each participant completing each subtest, a certified FMS tester instructed the participant in how to set up for each performance. Once set up was completed, the research assistant controlled test performance and instructed the participant when to start and finish each test performance. The certified FMS tester scored performance live using the recommended FMS score sheet in the FMS Manual (Appendix I) and remained blinded to prototype v3 scores during screening.

Each participant completed the deep squat three times, and the hurdle step and inline lunge subtests three times on each leg. A strict protocol was put in place by the researcher to ensure each subtest was conducted in a controlled manner and to ensure that the chance of bias was kept as low as possible. To ensure that prototype v3 had every opportunity to provide the correct score during the deep squat subtest, each participant completed one attempt without their heels on the board, hence giving them the opportunity to score a three, one or zero, and then two attempts with their heels on the board, giving them the opportunity to score a two, one or zero. Prototype v3 recorded all test performances so the best score from the three attempts was recorded following test completion. Although this process did not strictly follow the FMS Manual (Functional Movement Systems, 2017) instructions, as the quality of performance was less relevant for the purposes of the validation testing, it was felt that this was the most time efficient procedure that would reduce the chance of bias and provide prototype v3 with the opportunity to provide all four scoring outcomes. The exact protocol followed by the certified FMS tester and research assistant is shown in Figure 5.3. As the set up for the hurdle step and inline lunge subtests do not change depending on performance, the standard FMS Manual protocols were followed for these two subtests but were again led by the research assistant to keep the instructions consistent and clear for the participants.

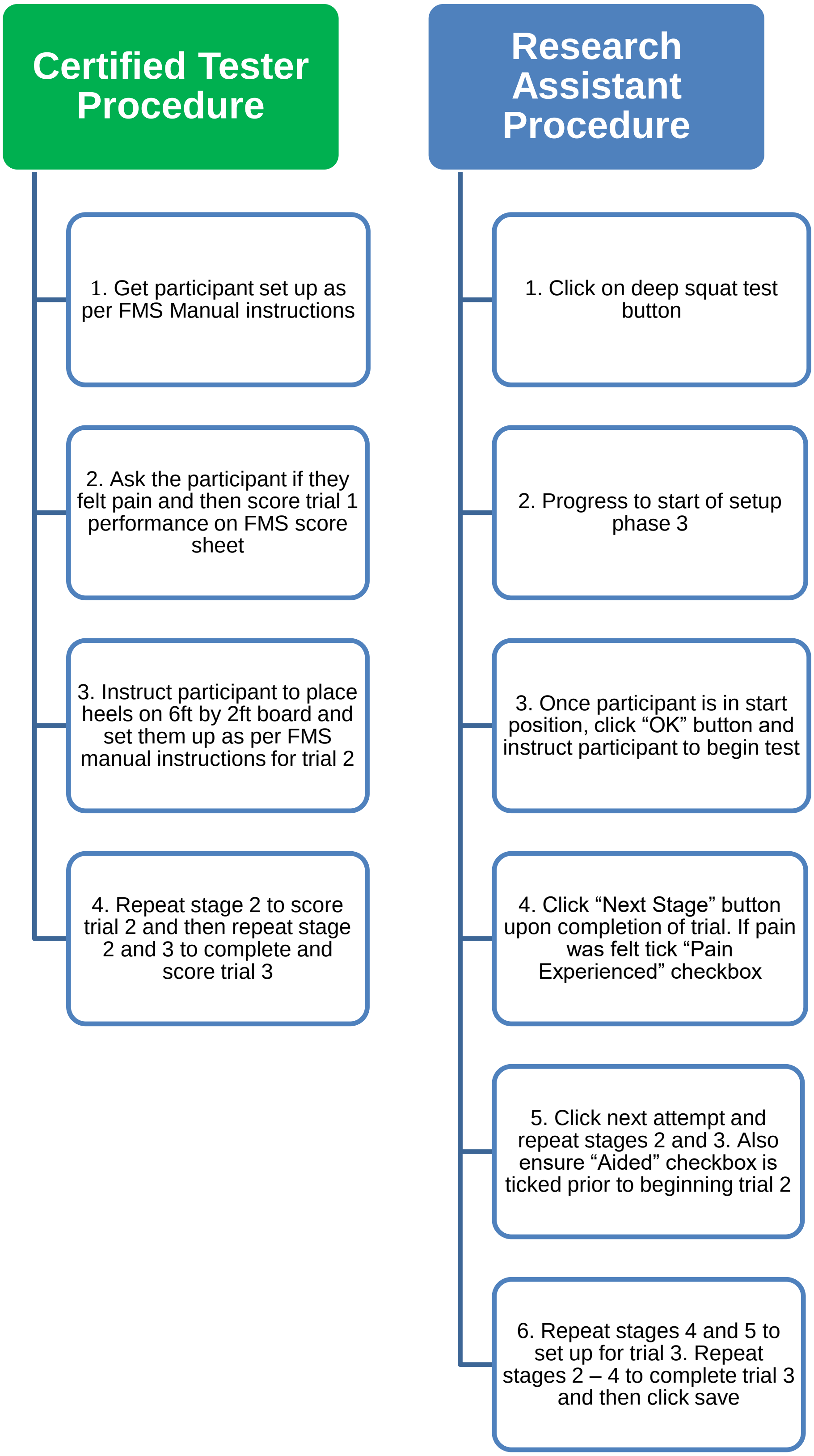


Figure 5.3. Protocol employed for validation of prototype v3 when assessing deep squat performance

Each participant was set up as a new user, under their participant code, within prototype v3's database and all scores automatically calculated by prototype v3 were stored following subtest performance. These scores were then manually recorded on to an Excel spreadsheet by the researcher. The manual scores recorded by the certified FMS tester were manually added to the same spreadsheet prior to the software scores being inserted.

Whilst prototype v3 was recording subtest performances, "Movavi" Screen Recording software was utilised to record the process. This allowed performances to be reviewed after the results had been analysed and statistical analysis was complete, to identify if any custom gestures for the three subtests were prone to identifying errors that were in fact not seen as errors by manual scoring. This process helped to form the basis for development of prototype v4 and the analysis that was completed using this process is discussed in chapter six.

5.3.3 STATISTICAL ANALYSIS

To ascertain if prototype v3 could accurately assess the FMS deep squat, hurdle step and inline lunge subtests compared to an expert tester, statistical tests were performed using IBM SPSS software version 24. Mean subtest scores across all participants were calculated for both prototype v3 and manual scoring methods. To assess the validity of prototype v3 as a method to assess subtest performance, the absolute and percentage agreement between both scoring methods for all participants scores were calculated. Additionally, a Spearman rank-order correlation test was conducted to understand the level of correlation between prototype v3 and manual scoring, as another method to test the validity of the prototype. To analyse inter-rater reliability between prototype v3 and manual scoring, a weighted kappa (K_w) statistic was calculated. A weighted kappa score of 0.80 and above indicates excellent agreement, 0.60 to 0.79 represents substantial agreement, 0.40 to 0.59, moderate agreement and below 0.40, poor to fair agreement. This analysis is in line with previous research (Teyhen et al, 2012; Shultz et al, 2013; Parenteau et al, 2014).

5.4 RESULTS (PROTOTYPE V3)

Twenty-seven participants performed the deep squat subtest three times and the hurdle step and inline lunge subtests three times on each leg. The mean subtest and total FMS composite scores provided by the manual and prototype v3 scoring methods are highlighted in Table 5.5.

Table 5.5. Mean subtest and total FMS composite scores provided by both manual and prototype v3 scoring methods for all participants (n=27)

FMS Subtest	Manual Tester Mean Scores	Prototype v3 Mean Scores
Deep Squat	1.2 (SD±0.4)	1.1 (SD±0.3)
Hurdle Step Right Side	2.2 (SD±0.4)	2.0 (SD±0.0)
Hurdle Step Left Side	2.0 (SD±0.2)	2.1 (SD±0.3)
Hurdle Step Total Score	2.0 (SD±0.0)	2.0 (SD±0.0)
Inline Lunge Right Side	2.2 (SD±0.5)	2.2 (SD±0.4)
Inline Lunge Left Side	2.0 (SD±0.2)	2.0 (SD±0.2)
Inline Lunge Total Score	2.0 (SD±0.3)	2.0 (SD±0.2)
Total FMS Composite Score	5.2 (SD±0.6)	5.1 (SD±0.4)

SD=Standard Deviation

When considering absolute agreement in scores between the two scoring methods, overall agreement for the deep squat subtest was 85%. Prototype v3 correctly identified 100% of performances scored as a one by manual scoring, but only identified two out of six performances as a two compared to manual scoring (Table 5.6).

Table 5.6. Breakdown of absolute agreement between prototype v3 and manual scoring for scores of one, two and three on the deep squat FMS subtest

	Manual Scores			
Prototype v3 Scores	**1**	**2**	**3**	**Total**
1	21	3	1	25
2	0	2	0	2
3	0	0	0	0
Total	21	5	1	27

For the hurdle step subtest, agreement for the total score was 100% between manual and prototype v3 scoring. Both scoring methods provided a total score of two for all participants. For unilateral performance, agreement for right side performance was 81%

and for left side performance was 96% between manual and prototype v3 scoring (Table 5.7).

Table 5.7. Breakdown of absolute agreement between prototype v3 and manual scoring for scores of one, two and three on the hurdle step FMS subtest

	Scores of 1	Scores of 2	Scores of 3	Agreement
Hurdle Step Prototype v3 Right Leg	0	27	0	81%
Hurdle Step Manual Right Leg	0	22	5	
Hurdle Step Prototype v3 Left Leg	0	26	1	96%
Hurdle Step Manual Left Leg	0	27	0	

For the inline lunge subtest, agreement for the total score was 96% between manual and prototype v3 scoring. For unilateral performance, agreement for right side performance was 93% and for left side performance was 100% between manual and prototype v3 scoring (Table 5.8).

Table 5.8. Breakdown of absolute agreement between prototype v3 and manual scoring for scores of one, two and three on the inline lunge FMS subtest

	Scores of 1	Scores of 2	Scores of 3	Agreement
Inline Lunge Prototype v3 Right Leg	0	23	4	93%
Inline Lunge Manual Right Leg	1	21	5	
Inline Lunge Prototype v3 Left Leg	0	26	1	100%
Inline Lunge Manual Left Leg	0	26	1	

For total subtest scores prototype v3 had a good, positive correlation with manual scoring for the deep squat (r_s=0.52; p=0.01), a very strong positive correlation with manual scoring for the hurdle step (r_s=1.00; p=0.01) and a strong positive correlation with manual scoring for the inline lunge (r_s=0.72: p=0.01).

Inter-rater reliability between manual and prototype v3 scoring for total subtest scores was substantial for the deep squat (K_w=0.63), perfect for the hurdle step (K_w=1.00), and substantial for the inline lunge (K_w=0.66) (Figure 5.4).

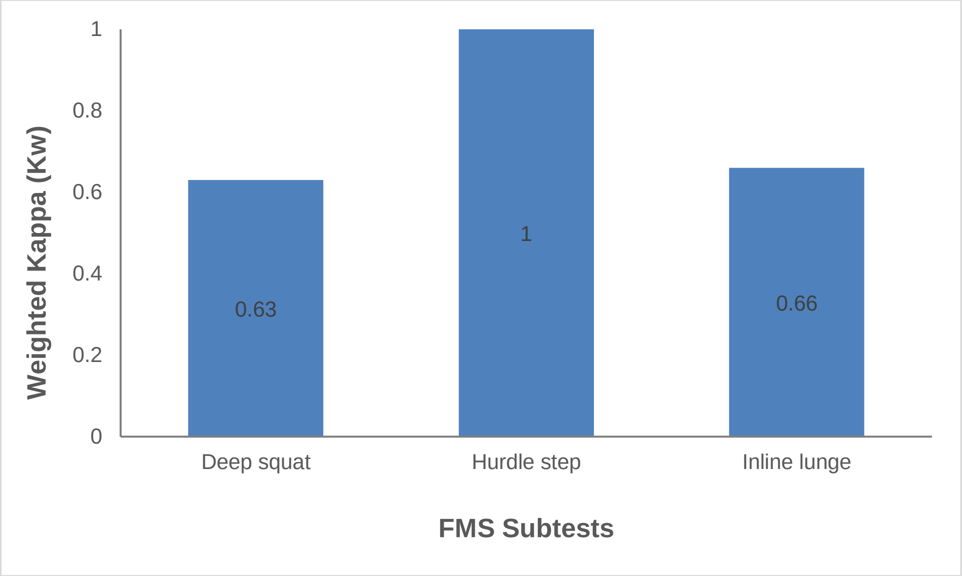


Figure 5.4. Inter-rater reliability values ($K_{w)}$ between prototype v3 and manual scoring for three FMS subtests

For unilateral assessment, inter-rater reliability was poor for the hurdle step right side (K_w=0.29), substantial for the hurdle step left side (K_w=0.65), substantial for the inline lunge right side (K_w=0.76) and perfect for the inline lunge left side (K_w=1.00) (Figure 5.5).

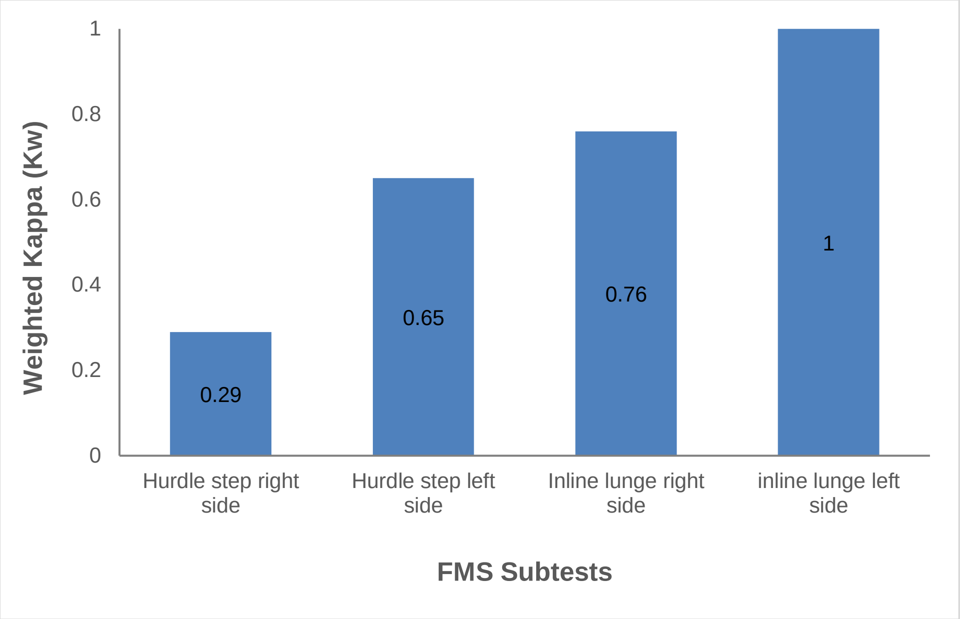


Figure 5.5. Inter-rater reliability values (K_{w}) between prototype v3 and manual scoring for unilateral hurdle step and inline lunge performance

5.5. DISCUSSION (PROTOTYPE V3)

The results indicated that prototype v3 had a good level of agreement with manual scoring when assessing performance on the three subtests. For the deep squat subtest the overall agreement of 85% between the two scoring methods was comparable to previous research in this area (Minick et al., 2010). However, the quality of performances should be considered, as over 75% of performances were scored as a one. As such it was not unexpected that the two scoring methods agreed, as a performance of one is possibly the easiest of the four scores to identify due to the obvious faults that occur when a participant was not able to score any higher. The results also highlighted that prototype v3 could not accurately identify performances of two compared to manual scoring. This suggested that the custom gestures set up to assess deep squat performance were still too sensitive and made it very difficult for a performance to be scored a two or above. A participant would need an almost perfect performance with minimal or no misalignment at key joints for prototype v3 to not detect a fault. Previous research, that developed a semi-automated system to assess the deep squat, was able to correctly identify 85.7% (n=7) of performances scored as a two by manual scoring, suggesting greater levels of accuracy than prototype v3 (Jensen, Weilbrenner, Rott, & Eskofier, 2013). Further analysis of the data from the validation testing completed for prototypes v1 to v3 was required to better understand the custom gestures causing this sensitivity and how they could be improved during further development work. This included reassessing all previous performances using new thresholds for each custom gesture to assess if the software correctly identified performances of a two or a three, as measured by manual scoring.

The agreement levels of the hurdle step and inline lunge subtests were high compared to manual scoring and showed promise for future development. The agreement for unilateral and total hurdle step score was higher than that seen in previous research comparing two or more manual raters (Parenteau et al., 2014). However, the similarity in performances across the sample could have influenced these results, with most participants achieving a score of two. On the hurdle step subtest, a score of two was the simplest performance to assess as there is broad scoring criteria available, meaning a relatively poor performance could still achieve a two, as could a very strong performance that did not quite reach the level of a three. As such, this level of agreement was not unexpected, but suggested further development work must focus on improving the prototype’s ability to accurately identify performances of three. Five performances on the right side were scored as a three by manual assessment but were not identified by prototype v3. This highlighted that the custom gestures were too sensitive, thus resulting in faults being identified that in fact did not exist according to manual scoring.

The inline lunge subtest showed the highest level of agreement with manual scoring across the three subtests, and again produced higher percentage agreements than previous research in this area (Minick et al., 2010; Parenteau et al., 2014). As per the other two subtests, there was similarity in performance levels across the sample, with many participants scoring a two, which could have influenced the agreement levels. However, unlike the hurdle step subtest assessment, prototype v3 was able to correctly identify all but one performance of a three. This suggested that the custom gestures for the inline lunge subtest were set at an appropriate level of sensitivity to accurately track performance. It should be noted that one performance that was manually scored as a one due to a loss of balance could not be identified as a one by prototype v3 as the “Loss of Balance” check box did not work effectively. This needed to be corrected in future development work.

The levels of concurrent validity reported for total scores suggest that prototype v3 could validly be used to assess hurdle step and inline lunge performance but was still not at an acceptable level for deep squat assessment. Prototype v3 reported perfect correlation for the hurdle step and strong correlation for the inline lunge compared with manual scoring, yet only good correlation for the deep squat. Although levels of correlation suggest a relationship does exist between the two methods, they do not necessarily identify true levels of agreement, and as such should be considered with caution in this instance. It is also clear from the results that the level of correlation for the deep squat subtest was affected by the fact prototype v3 was not able to identify certain scores of two or three. The deep squat was the only subtest where more than two different total scores were provided for all participants by manual scoring, hence affecting levels of correlation compared to those subtests that only reported one or two different total scores across the sample. The number of correctly identified deep squat performances by prototype v3 was comparable to those for the other two subtests, but the incorrect scores caused the levels of correlation to be reduced.

Inter-rater reliability results show that the prototype v3 was a reliable method to assess the three FMS subtests compared to manual assessment, with a higher K_w reported for total subtest scores compared to previous research (Shultz et al., 2013; Teyhen et al., 2012b). However, this level of reliability reduces when considering unilateral performance for the hurdle step subtest and increases for unilateral performance on the inline lunge subtest. Levels of reliability for unilateral performance were similar in the hurdle step and higher in the inline lunge subtest to previous research that developed an automated system to assess six FMS subtests (Whiteside et al., 2014). This previous research also showed lower levels of agreement between manual scoring and the automated scoring system for total scores on both subtests, compared to prototype v3, suggesting unilateral performance heavily influenced the total scores provided by the automated IMU system in the study. In the current study, the level of reduction in unilateral agreement for the hurdle step subtest was affected by the fact prototype v3 failed to identify five performances on the right side as a three compared to manual scoring. However, these unilateral performances did not impact the total score given by prototype v3, as all but one of them scored a two overall, but these performances unduly effected the K_w values for the right-side assessments. The hurdle step subtest involves movement of one limb (the right or left leg) whilst the rest of the body is kept as still as possible. Previous research has highlighted that the Kinect can have difficulty when tracking lower limb movement during different activities (Bonnechère et al., 2014) and therefore prototype v3 may have had difficulty accurately tracking right or left leg movement. This was analysed during development of prototype v4 to fully understand if the active leg was accurately tracked. For the inline lunge subtest, K_w values increased for unilateral assessment. For right side assessment, only two performances were incorrectly scored by prototype v3, and for the left side, all performances were scored correctly, leading to high inter-rater reliability values and suggesting that the positioning of the FMS board at 30° to the Kinect sensor helped enable prototype v3 to accurately assess performance.

5.6 CONCLUSION (PROTOTYPE V3)

This study provided an overview of the development of prototype v3 to automatically assess three FMS subtests, as well as discussing the validation process undertaken to assess its validity and reliability compared to manual FMS assessment. The development process involved adjusting the existing software that had been incorporated into prototype v1 and v2 to assess deep squat performance. These adjustments were designed to improve prototype v3's ability to accurately assess deep squat performance, and to allow prototype v3 to assess hurdle step and inline lunge subtest performance. A similar process was adopted for these two subtests, in regard to the custom gestures employed to automatically assess performance, by following the scoring criteria for each subtest as set out in the FMS Manual. Although it was not possible to incorporate machine learning into the assessment of these two subtests, the researcher and development team were still able to use joint angles and alignment as the key parameters against which performance was measured to produce a score of zero, one, two or three. Adjustments were also made to the user interface and software database, in attempt to improve efficiency and data storage capabilities, to maximise prototype v3's usability in applied settings.

Although the validity results for prototype v3 were very encouraging, improvements were still needed before the prototype could be used repeatedly in an applied setting to assess all seven FMS subtests. Further work was still required to adjust the custom gestures set for all three subtests to reduce the sensitivity of the prototype v3 and improve identification of scores of three, particularly for deep squat performance. Analysis of the prototype v3's screenshots from this validation helped to identify which custom gestures required further adjustment to improve these existing subtests and helped set up the remaining four subtests as accurately as possible for prototype v4.

Another area of improvement could be seen by changing the height of the Kinect sensor. For this study the sensor was set at 1.00 m high. Recent research suggested that placing the Kinect at a different angle and at a greater height may offer more accurate tracking of movement (Muller et al., 2017). With three of the remaining four FMS subtests involving the participant being in a prone or supine position, placing the sensor at a greater height for prototype v4, could allow the sensor to track more joints and limbs on each participant, which could lead to improved accuracy. The inline lunge subtest results showed that adjusting the set up of the Kinect sensor can have a positive influence over the scoring accuracy of the prototype.

Finally, future validation was needed with more than one certified FMS tester, and with assessment being completed with both video and live manual scoring, at different timepoints. The validation testing completed in this study only included assessment by one manual certified tester at one timepoint. To determine prototype v4’s validity, comparing its scores to more than one certified FMS tester was necessary. Using two or more testers would provide a more objective measure against which to compare the prototype, as it would remove the possibility of any scoring bias from one tester. Additionally, analysing the agreement between prototype v4 and both live and video scoring, would provide another indication of its validity and reliability against the current gold standard methods of assessment. Although previous research has highlighted that manual scoring may not be a valid reference against which to compare automated systems (Whiteside et al., 2014), due to the subjective nature of the scoring system and each testers interpretation of same, it was the only standard available to assess the prototypes against in this programme of research. In addition to comparing prototype v4’s validity to live and video manual scoring, it was also important for the test-retest reliability of prototype v4 to be determined to ensure it was consistent in the scores that it provided. This study only assessed prototype v3 scoring at one particular point in time. For prototype v4 to be an effective tool in applied settings, it required the ability to consistently provide the same scores across different testing sessions.

CHAPTER SIX

DEVELOPMENT OF MOTION TRACKING SOFTWARE TO ASSESS THE SEVEN FMS SUBTESTS

6.1 INTRODUCTION

The previous study reported that prototype v3 had good agreement compared to manual scoring of performances for three FMS subtests However, further work was required before the software could be used in an applied setting and be used as a truly valid alternative to manual assessment. When considering the three subtests that prototype v3 assessed, amendments to the custom gestures were required to help improve the accuracy of the software's scoring. In particular, the deep squat subtest required improvement in its ability to score performances of two or three. It had high levels of agreement when scoring performances of one, but often identified faults that would not be classified as a fault by manual scoring. In addition, further assessment of the hurdle step subtest was required to analyse the prototype's ability to track active leg movement and to understand whether adjustments to the custom gestures were required to improve accuracy in this area.

The FMS is made up of seven subtests, and prototype v3 was only able to assess three of these. Therefore, development of prototype v4 focussed on the additional four subtests and how it could best be designed to automatically assess performance on these subtests. A full description of each FMS subtest is available in Appendix F. Consideration was given to the machine learning capabilities of the Kinect sensor and whether these could be incorporated into the custom gestures for each subtest. In addition, the height that the Kinect sensor was set at was adjusted to maximise the chances of accurately tracking joint centres throughout performance on subtests that involved a participant lying in a prone position (trunk stability), supine position (active straight leg raise) and quadruped position (rotary stability). Increasing the height from 1.00 m provided a better opportunity for accurate joint tracking. This adjusted height was also incorporated into the existing subtests that prototype v3 assessed, to reduce time and improve efficiency when using the software in applied settings.

As discussed in previous chapters, the FMS Manual (Functional Movement Systems, 2017) provides an algorithm that describes how to interpret subtest scores to best inform the prescription of corrective exercise interventions. These interventions should focus on weaknesses within mobility subtests first (ASLR and SM) followed by weaknesses in stability subtests (RS and TS) and then finally any weaknesses identified on the remaining three total body movement subtests. It is therefore imperative that the subtests are scored with accuracy and precision to ensure the correct interventions are prescribed by practitioners to each individual. As part of the development of prototype v4 a corrective exercise prescription function was added, that automatically prescribed an intervention for an individual based on their FMS subtest results. This automatic system was based on the corrective exercise algorithm specified in the FMS manual to ensure consistency with manual methods.

The test-retest reliability and validity of prototype v4 was measured against gold standard methods to assess if it met acceptable levels of agreement with these methods and could validly be used in applied settings. Previous research has shown that video recording of the FMS is a reliable method to assess performance (Minick et al., 2010; Parenteau et al., 2014), yet using video scoring does not truly replicate assessment of the FMS in applied settings (Smith et al., 2013). Therefore, for the purposes of validating prototype v4 and assessing its validity against manual scoring, both live and video scoring were incorporated into this study, which provided a more accurate reflection of the prototype's applicability in applied settings. Two certified FMS testers were employed to assess performances, one through live manual scoring and the other through video assessment conducted post performance.

6.2 RESEARCH AIMS (PROTOTYPE V4)

This study attempted to answer the following research questions:

- Could software, integrated with the Kinect sensor v2, be developed that could validly assess performance on the seven FMS subtests compared to blinded manual screening by two certified FMS testers?
- Could software, integrated with the Kinect sensor v2, be developed that increased the test-retest reliability of the seven FMS subtests?
- Could software, integrated with the Kinect sensor v2, be developed that could accurately prescribe corrective exercises based on FMS subtest scores, compared to two certified FMS testers?

6.3 METHODS (PROTOTYPE V4 DEVELOPMENT)

6.3.1 PROTOTYPE V4 DEVELOPMENT PHASE

Prior to beginning development on prototype v4, a detailed plan was developed that broke the work down into specific units of work based upon previous prototype development and any new functionality required for prototype v4 (Table 6.1).

Table 6.1. Development plan for prototype v4

Task	Effort
Raw Data Review for remaining 4 Subtests	3 Weeks
Machine Learning Algorithm for remaining 4 Subtests	2 Weeks
Custom Gesture Development/Adjustment	4 Weeks
Corrective Exercise Functionality	4 Weeks
Initial Prototype Testing	2 Weeks
Add Initial Test Scoring	2 Weeks
Testing & Debugging	2 weeks
Total	**19 Weeks**

Raw data review

The raw data captured prior to prototype v2 development (see Chapter 4) was used to assess if machine learning could be incorporated into any of the remaining four subtest custom gestures.

Machine learning

The raw data highlighted that applying machine learning to the remaining subtests would have been a complicated process that may not have improved the automated scoring system, and from a time perspective could not be completed within the scope of this research. For the active straight leg raise, it was found that machine learning would have had limited relevance as the score achieved was directly related to the mid-thigh measurement which differed for each individual. Hence, within the time available for this study it was decided to build a custom gestures template that could accurately replicate a score of three, two or one for every participant rather than use the machine learning approach. For the trunk stability and rotary stability subtests, upon analysis of the video footage, building custom gestures based around joint angles and alignment would be a more accurate method to use compared to a machine learning template for each possible score. The two subtests involved different set up positions for scores of two and three, and the finish points of each were dependent on the length of each individual's limbs. This made it difficult to build a template related to subtest progression as per the deep squat. The shoulder mobility subtest did not utilise the Kinect sensor's joint tracking or machine learning capabilities. Upon review of the raw data, it was found that the Kinect could not accurately track the hand positions once they were placed behind the back, and therefore a different approach was used for this subtest which is outlined later in the chapter.

Custom gesture development

Chapters four and five discussed the progressions that were made from development of prototype v1 through to prototype v3, and how the accuracy of the automated software improved through the analysis of each validation to better understand how and why faults had been detected. This analysis allowed small adjustments to the custom gestures for the deep squat subtest, and to build a custom gesture table for the hurdle step and inline lunge subtests, based on the FMS Manual scoring criteria. The same approach was adopted for the development of prototype v4, whereby the validation of prototype v3 was assessed to improve the accuracy of the three existing subtests. In addition, the FMS Manual's scoring criteria for the four remaining subtests were used as the basis for the custom gesture tables developed for each.

Deep Squat

As outlined in chapter five, the reliability and validity of prototype v3 when assessing the deep squat subtest required improvement. Prototype v3 was inaccurate when scoring performances of two or three, and often scored a performance too low due to identifying a fault that was not identified as a fault by manual scoring. The four performances from prototype v3 validation which the prototype scored as a one, but manual assessment scored as a two, were analysed and it was identified that two of the custom gestures generated faults for all four performances. These were the "Squat Progress" and "Knee inline with Ankle" gestures. The "Squat Progress" was based on machine learning and had remained at a setting of 100% throughout prototype development. All four performances from prototype v3 validation failed to reach the 100% mark according to the software yet had managed to get their femur below horizontal. All performances achieved a progress percentage of between 95% and 99% according to the software, meaning prototype v3 identified a fault, as the measurement set for the gesture was incorrect. For prototype v4 the measurement for this custom gesture was changed to 95%.

The “Knee inline with Ankle” gesture identified a fault for three of the four participants, suggesting that there had been lateral movement of the knees resulting in them not being in line with their ankles. As manual scoring had not identified this as a fault, this gesture was still too sensitive and incorrectly detecting faults. As this gesture also caused issues for prototype v2, the gesture and the movement that it measured were changed. For prototype v4 this gesture was changed to “Distance between Knees” and assessed the distance on the x-axis between the two knees. Prior to performance the prototype measured the distance between the knees and tracked whether this changed during performance, allowing movement of 0.03 m each way. If the distance changed beyond 0.03 m in either direction, the gesture identified a fault. Table 6.2 provides details of the final custom gestures set for prototype v4 to assess the deep squat.

Table 6.2. Custom gestures used to identify movement errors in prototype v4 during deep squat subtest performance

Deep Squat subtest	Description	Units	Min	Max
1. Set Up Phase 1				
Bar on Head	Checks the angle of each shoulder and bar on head	Degrees	165.00	175.00
Hands above Head	Measures distance between hand and head on the Y-axis	Metres	0.00	0.10
2. Set Up Phase 2				
Hands to Head Distance	Measures distance between hand and head on the X-axis	Metres	0.00	0.10
Hands Behind Head	Measures distance between the hand and foot on the Z-axis	Metres	0.00	0.10
3. Test Phase				
Squat Progress	Using machine learnt code, determines the progress of the user in the squat	Percentage	N/A	95.00
Hands behind Head	Measures distance between the hand and foot on the Z-axis	Metres	0.00	0.10
Foot inline with Ankle	Calculates the distance on the X-axis between the ankle and foot	Metres	-0.02	3.00
Parallel Trunk/Tibia	Compares angle of the spine in YZ plane and the angle of the tibia in YZ plane	Degrees	30.00	N/A
Distance between Knees	Calculates the distance on the X-axis between each knee	Metres	-0.03	0.03

Hurdle Step

The same process was adopted for the hurdle step subtest to assess if any custom gestures were frequently identifying faults and causing performance to be measured a two rather than a three. Although the agreement levels with manual scoring was high for the prototype v3, for unilateral performance there were several performances incorrectly scored as a two by the prototype instead of a three on the right side. When reviewing the results, two custom gestures were overly sensitive. The “Hip extension” and “Active Leg Knee Alignment” gestures both identified faults that manual scoring did not, and as such the range for both gestures was extended for prototype v4 by 2° for “Hip Extension” and 0.02 m for “Active Leg Knee Alignment” (Table 6.3). No other adjustments were made to the custom gestures for this subtest.

Table 6.3. Custom gestures used to identify movement errors in prototype v4 during hurdle step subtest performance

Hurdle Step subtest	Description	Units	Min	Max
1. Set Up Phase 1				
Hand to Neck Distance	Compares distance between hand and neck on each side	Metres	0.00	0.01
2. Set Up Phase 2				
Hand to Neck Distance	Compares distance between hand and neck on each side	Metres	0.00	0.01
3. Test Phase				
Hip Extension	Measures the angle of the hip on YZ Plane	Degrees	85.00	95.00
Active Leg Knee Alignment	Measures the distance on X-axis of knee and hip on active leg	Metres	-0.02	0.05
Active Leg Ankle Alignment	Checks distance on X-axis of ankle and knee on active leg	Metres	0.00	0.05
Lateral Movement	Measures the angle of the spine on X-axis	Degrees	85.00	95.00
Upper Body Rotation	Compares distances on the Z-axis of shoulder to hip on each side	Metres	0.00	0.05

Inline Lunge

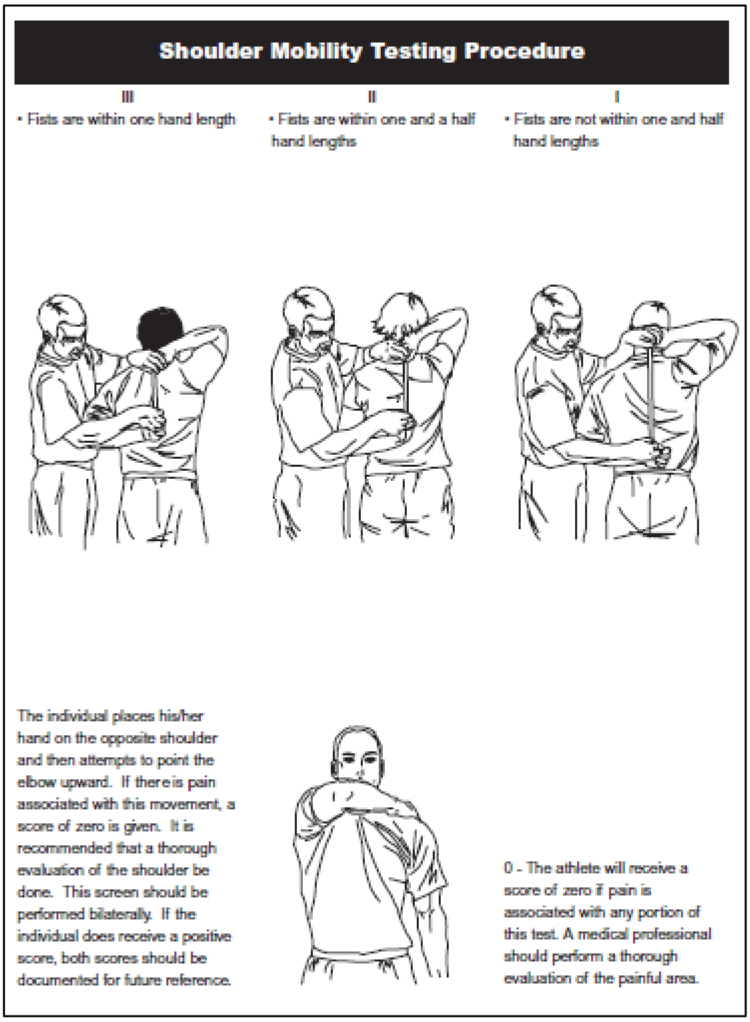
The inline lunge subtest showed good levels of agreement and reliability compared to manual scoring for prototype v3. There were only two unilateral performances on the right side scored incorrectly by the prototype v3. On review of prototype v3 results, one gesture was unable to identify a fault for one performance that should have been scored a one but was scored as a two. The range of the “Body Lean” gesture appeared to be too wide and was adjusted for the development of prototype v4 by 3° as per Table 6.4. All other custom gestures remained as per prototype v3.

Table 6.4. Custom gestures used to identify movement errors in prototype v4 during inline lunge subtest performance

Inline Lunge subtest	**Description**	**Unit**	**Min**	**Max**
1. Test Phase				
Body Lean	Checks torso angle	Degrees	88.00	95.00
Heel Flat	Checks front heel remains flat on the board	Metres	0.00	0.01
Knee and Ankle Inline	Checks distance of knee and ankle on the Z-axis	Metres	0.00	0.05
Lateral Movement	Measures the angle of the spine on X-axis	Degrees	85.00	95.00
Upper Body Rotation	Compares distance between shoulder and hip on the X-axis	Metres	0.00	0.02

Shoulder Mobility

The shoulder mobility subtest assesses an individual's bilateral shoulder range of motion in the shoulder joints (see Appendix F for full description). The scoring criteria for the shoulder mobility subtest and instructions for the clearing test that are described in the FMS Manual are outlined in Figure 6.1.



Shoulder Mobility Testing Procedure

III
- Fists are within one hand length

II
- Fists are within one and a half hand lengths

I
- Fists are not within one and half hand lengths

The individual places his/her hand on the opposite shoulder and then attempts to point the elbow upward. If there is pain associated with this movement, a score of zero is given. It is recommended that a thorough evaluation of the shoulder be done. This screen should be performed bilaterally. If the individual does receive a positive score, both scores should be documented for future reference.

0 - The athlete will receive a score of zero if pain is associated with any portion of this test. A medical professional should perform a thorough evaluation of the painful area.

Figure 6.1. The shoulder mobility FMS subtest scoring criteria and clearing test instructions (Functional Movement Systems, 2017)

As shown in Figure 6.1, the scoring criteria for the shoulder mobility subtest is relatively simple and relates to the distances between the participant's fists:

- A score of three is achieved if the fists are within one hands length
- A score of two is achieved if the fists are within one and half hand lengths
- A score of one is given if the fists are not within one and half hand lengths
- A score of zero is given if the individual reports pain during the clearing test

After initial testing and raw data analysis, the Kinect sensor had difficulty tracking the hands and arms once they were placed behind the participants back during subtest performance. The joint centres were unable to be accurately detected and it was not possible to create custom gestures for this subtest. Instead, a semi-automated system was developed for prototype v4 to assess shoulder mobility performance (Figure 6.2).

1. Length of participants hand was measured by tester. Measurement was inputted into user interface

2. Participant performed subtest and distance between the fists was measured by tester as per FMS Manual guidelines. Measurement inputted into user interface

3. Prototype v4 used both measurements to calculate whether fists were within one hands length, one and a half hand lengths or wider. Score was automatically provided based on this calculation

4. Test was repeated upto three times on each side by following steps 1 to 3. Clearing test was completed by tester before final score was provided by prototype v4

Figure 6.2. Prototype v4's semi-automated scoring system for the shoulder mobility subtest

Active Straight Leg Raise

This subtest is designed to assess hip mobility in the passive leg, whilst measuring hamstring and gastric-soleus flexibility in the active leg. Three performances on each leg are attempted and scored as per the scoring criteria set in Figure 6.3.

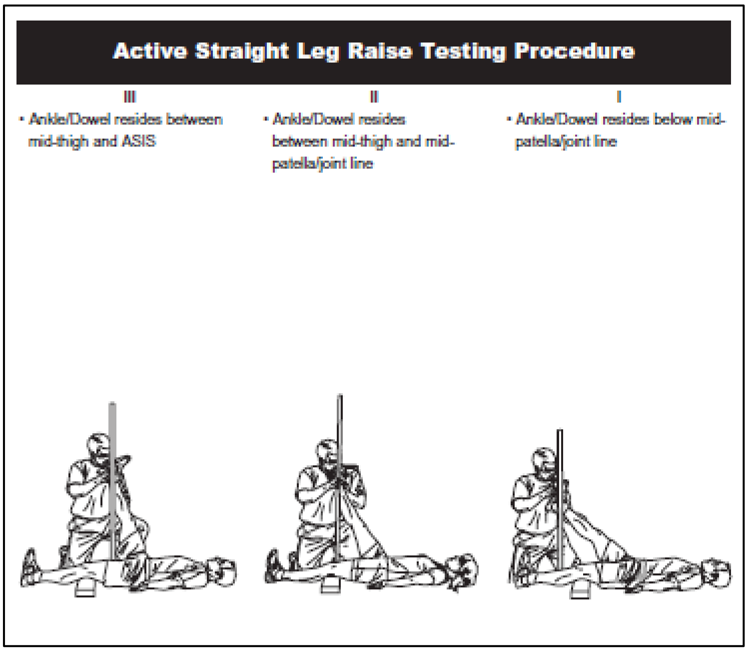


Figure 6.3. The active straight leg raise FMS subtest scoring criteria (Functional Movement Systems, 2017)

The scoring criteria relates specifically to the point at which the malleolus on the active leg resides at end range.

- Extending the malleolus past the mid-thigh point is classified as a score of three.
- Extending the malleolus between mid-thigh and mid patella is classified as a score of two
- Failure to extend the malleolus past the mid patella point is classified as a score of one

For the purposes of prototype v4, a custom gesture table was developed that encompassed the key criteria outlined above to allow automated assessment. Table 6.5 provides details of the custom gestures developed for the active straight leg raise.

Table 6.5. Custom gestures used to identify movement errors in prototype v4 during active straight leg raise subtest performance. Gestures highlighted in yellow were removed prior to validation

Active Straight Leg Raise	Description	Units	Min	Max
1. Set Up Phase				
Hand Movement	Compares distance between hand and shoulder on the X-axis.	Metres	0.00	0.01
2. Test Phase				
Hand Movement	Compares distance between hand and shoulder on the X-axis.	Metres	0.00	0.01
Passive Leg	Checks for any movement in the passive leg	Metres	0.00	0.01
Active Leg Bend	Checks angle of active leg and checks for any bend in it	Degrees	177.00	180.00
Active Leg Raise Angle	Displays angle of active leg that is being raised	Degrees	N/A	N/A
Ankle not passing Knee	Checks the distance between the ankle and the passive knee. Takes in an inputted value to assist in calculating the scores	Metres	N/A	N/A

Initially the custom gestures for the active straight leg raise were split into two distinct phases as highlighted in Table 6.5.

1. Set up phase

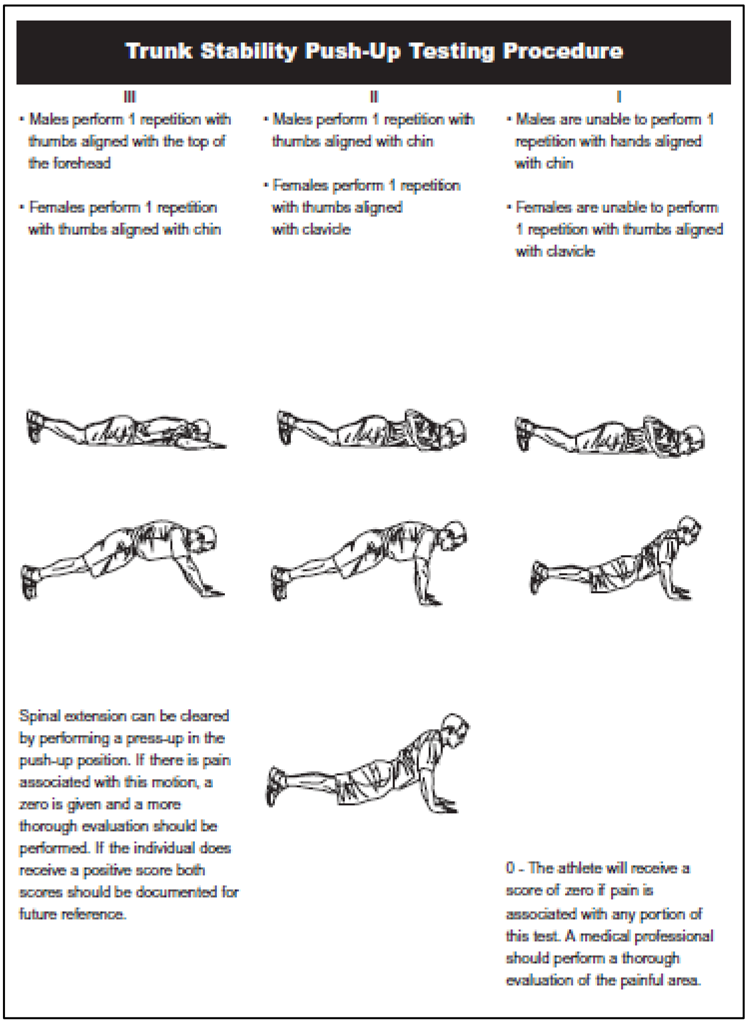
The "Set Up Phase" was designed so prototype v4 could detect whether the participant was correctly in position prior to performance by measuring the distance between the hands and shoulders. However, after conducting pilot testing using these measures, the Kinect sensor was not able to provide the level of accuracy required when tracking the hands and shoulders furthest from the camera due to the participant being in the supine position. Therefore the "Set Up" phase was removed from the custom gesture table for this subtest.

2. Test phase

The "Test Phase" initially included five different custom gestures to assess test performance. The "Hand Movement" gesture had the same difficulty as the same gesture in the "Set Up Phase" and was removed for the same reason. The "Passive Leg" gesture was designed to track movement in the passive leg during test performance. However, as per the "Hand Movement" the Kinect sensor was unable to accurately track any movement that occurred in the key joints within the passive leg, due to it being placed furthest from the sensor during performance. This caused the prototype v4 to incorrectly score performances during pilot testing and this gesture was removed. To replace these two automated gestures, a manual "Passive Movement" button was inserted into the user interface, which once manually pressed, used the current score at that point to score performance. The three remaining custom gestures related to the active leg. The “Active Leg Bend” measured whether the active leg flexed at the knee whilst it was being raised. It indicated a fault if the angle of extension was less than 177°, and automatically took the score at this point as the performance score. Prior to test performance, the distance between the ASLR and mid patella was entered into the user interface, and the prototype calculated the mid-thigh point using this information. The “Active Leg Raise Angle” and “Ankle not passing Knee” gestures were then used to detect the position of the malleolus on the active leg compared to the knee and calculated mid-thigh point on the passive leg, to provide a score for that performance. As per all other subtests, the tester had the option to manually check the “Pain” box on the user interface, which indicated pain was felt during performance and a score of zero was given by prototype v4.

Trunk Stability

The trunk stability subtest is designed to test an individual’s ability to stabilise the spine in an anterior and posterior plane during a close-chain upper body movement. The scoring criteria provided in the FMS Manual and clearing test required to be undertaken upon completion of the subtest are outlined in Figure 6.4.



Trunk Stability Push-Up Testing Procedure

III	II	I
• Males perform 1 repetition with thumbs aligned with the top of the forehead • Females perform 1 repetition with thumbs aligned with chin	• Males perform 1 repetition with thumbs aligned with chin • Females perform 1 repetition with thumbs aligned with clavicle	• Males are unable to perform 1 repetition with hands aligned with chin • Females are unable to perform 1 repetition with thumbs aligned with clavicle

Spinal extension can be cleared by performing a press-up in the push-up position. If there is pain associated with this motion, a zero is given and a more thorough evaluation should be performed. If the individual does receive a positive score both scores should be documented for future reference.

0 - The athlete will receive a score of zero if pain is associated with any portion of this test. A medical professional should perform a thorough evaluation of the painful area.

Figure 6.4. The trunk stability FMS subtest scoring criteria and clearing test instructions (Functional Movement Systems, 2017)

The scoring criteria for the trunk stability subtest are directly related to the hand position at the start of the test:

- For males, a score of three is given if a push up is performed with the thumbs positioned level with the top of the forehead
- For females, a score of three is given if a push up is performed with the thumbs positioned level with the chin
- For males, a score of two is given if a push up is performed with the thumbs positioned level with the chin
- For females, a score of two is given if a push up is performed with the thumbs positioned level with the clavicle
- A score of one is given to both males and females if they cannot perform a push up with their hands in the lower position.

Despite the different start positions for males and females, custom gestures were designed for prototype v4, to assess the set-up position and test performance for the trunk stability subtest, using the key criteria set in the FMS Manual (Table 6.6).

Table 6.6. Custom gestures used to identify movement errors in prototype v4 during trunk stability subtest performance

Trunk Stability	Description	Units	Min	Max
1. Set Up Phase				
Ankle/Knee Inline	Ensures both ankles are inline on the X-axis	Metres	0.00	0.05
	Ensures both knees are inline on the X-axis	Metres	0.00	0.05
Hands/Foot Inline	Ensures both hands are inline on the X-axis	Metres	0.00	0.05
	Ensures both feet are inline on the X-axis	Metres	0.00	0.05
2. Test Phase				
Hip Inline/Angle	Ensures hips are inline on the X-axis	Metres	0.00	0.03
	Checks the angle between left and right hips on Y-axis	Degrees	177.00	180.00
Shoulder Inline/Angle	Ensures both shoulders are inline on the X-axis	Metres	0.00	0.03
	Checks the angle between both shoulder joints on Y-axis	Degrees	177.00	180.00
Torso Inline/Angle	Ensures hip is inline with torso on the X-axis	Metres	0.00	0.03
	Checks the angle between hip and shoulder joints on Y-axis	Degrees	175.00	180.00

1. Set up phase

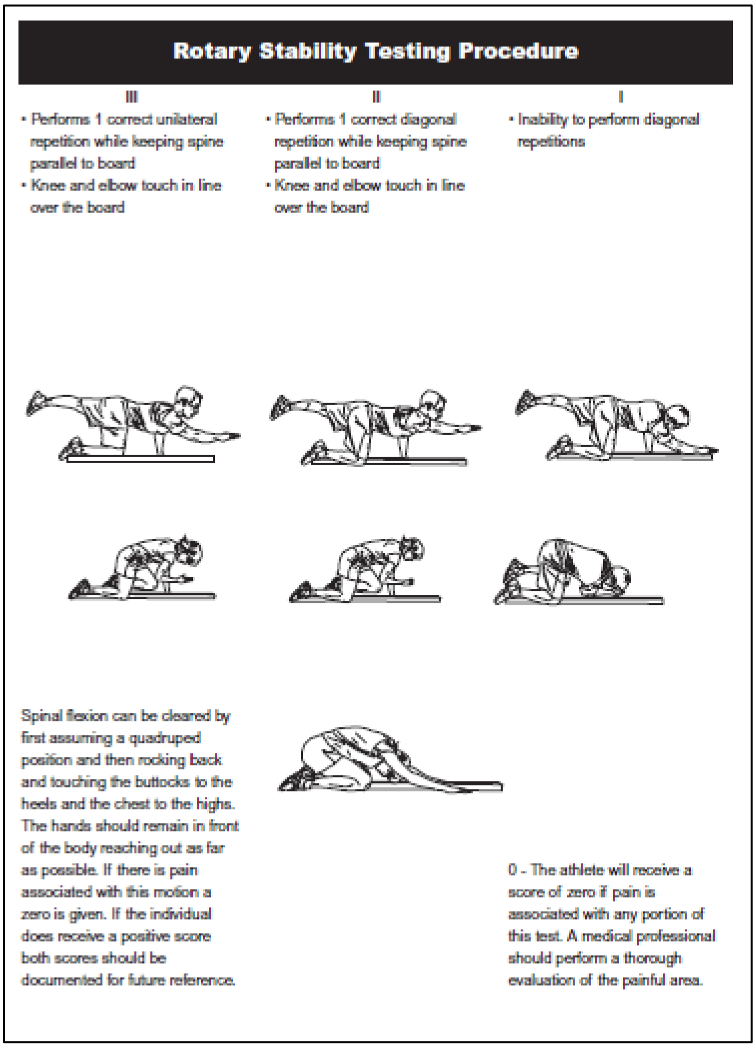
The position of the hands change depending on the gender of the participant and the score that is being attempted. As such assessment of this set up process was not automatically conducted by prototype v4. The positioning of the hands for each attempt was monitored by the tester and was left as a manual process prior to each performance. The four custom gestures that make up the “Set Up Phase” assessed whether the hands and feet moved on the x-axis. The gestures identified if any movement had occurred, so the tester could reset the set-up position prior to performance.

2. Test phase

The three custom gestures included in the “Test Phase” assessed movement in the trunk to check that the body moved as one unit during performance. The “Hip Inline/Angle” gesture checked the distance between the left and right hip on the x-axis didn’t move by more than 0.03 m and that the angle between the two joints on the y-axis didn’t change by 3° or more. Any tilting or sway within the torso causes the hips to move and led to a fault being identified by this gesture. The “Shoulder Inline/Angle” gesture checked for the same movement as the previous gesture but was focussed on the shoulder joints. This detected if one side of the body was pushed up before the other side, leading to a tilt in the shoulder area. The “Torso Inline/Angle” gesture, linked the top and bottom of the torso to check for lag in the lumbar spine region. It checked for a change in distance of 0.03 m or more between the hips and shoulders on the x-axis, whilst also checking for any change of 5° or more in the angle between the two on the y-axis. As per all other subtests, the tester has the option to manually check the “Pain” box on the user interface, which indicated pain was felt during performance and a score of zero was given by the prototype v4. More detail regarding how prototype v4 distinguishes between different set up positions to accurately score performance. is included in Appendix H.

Rotary Stability

This subtest assesses multi-plane trunk stability during a combined upper and lower extremity movement. The scoring criteria provided by the FMS manual and clearing test instructions are outlined in Figure 6.5.



Rotary Stability Testing Procedure

III	II	I
• Performs 1 correct unilateral repetition while keeping spine parallel to board • Knee and elbow touch in line over the board	• Performs 1 correct diagonal repetition while keeping spine parallel to board • Knee and elbow touch in line over the board	• Inability to perform diagonal repetitions

Spinal flexion can be cleared by first assuming a quadruped position and then rocking back and touching the buttocks to the heels and the chest to the highs. The hands should remain in front of the body reaching out as far as possible. If there is pain associated with this motion a zero is given. If the individual does receive a positive score both scores should be documented for future reference.

0 - The athlete will receive a score of zero if pain is associated with any portion of this test. A medical professional should perform a thorough evaluation of the painful area.

Figure 6.5. The rotary stability FMS subtest scoring criteria and clearing test instructions (Functional Movement Systems, 2017)

The scoring criteria highlighted in Figure 6.5. show that the score provided is directly related to the type of movement performed:

- A three is achieved if the spine remains parallel to the board during unilateral performance, and the knee and elbow touch over the board
- A two is achieved if the spine remains parallel to the board during diagonal performance, and the knee and elbow touch over the board
- A one is scored if the participant is unable to performance a diagonal repetition

For the purposes of using prototype v4 to automatically score performance, a “Set Up Phase” and “Test Phase” were included in the custom gesture table (Table 5.11).

However, upon completion of pilot testing certain custom gestures were removed, as highlighted in Table 6.7.

Table 6.7. Custom gestures used to identify movement errors in prototype v4 during rotary stability subtest performance. Custom gestures highlighted in yellow were removed prior to validation

Rotary Stability	Description	Units	Min	Max
1. Set Up Phase				
Hands Inline	Checks to ensure the hands are inline on the X-axis	Metres	0.00	0.03
Shoulders Inline	Checks to ensure the shoulder joints are inline on the X-axis	Metres	0.00	0.03
Knees Inline	Checks to ensure the knee joints are inline on the X-axis	Metres	0.00	0.03
Ankle/Foot Inline	Checks to ensure the ankle joints are inline on the X-axis	Metres	0.00	0.03
	Checks to ensure the feet are inline on the X-axis	Metres	0.00	0.03
2. Test Phase				
Passive Leg	Ensures active leg (right) is straight	Degrees	177.00	180.00
	Checks to ensure passive leg (left) doesn’t move on Y-axis	Metres	0.00	0.05
Passive Arm	Ensures active arm (left) is straight	Degrees	177.00	180.00
	Checks to ensure passive arm (right) doesn’t move on Y-axis	Metres	0.00	0.05
Spine Angle	Checks the angle of the torso on the Y-axis	Degrees	175.00	180.00
Shoulder Angle	Checks the angle between both shoulder joints on the X-axis	Degrees	177.00	180.00
Hip Angle	Checks the angle between both hip joints on the Z-axis	Degrees	175.00	180.00

Set up phase

As per the trunk stability subtest, the set-up phase used gestures to check the alignment of a variety of joints on the x-axis, after a participant has been correctly set up manually by the tester. Any movement of the hands, feet, knees, shoulders and ankles once in this position were identified as a fault by prototype v4, allowing the tester to reset the participants position.

Test phase

Initially, five custom gestures were built for the rotary stability subtest to assess performance. However, upon completion of pilot testing, amendments were made to two of these gestures and one was removed. The “Passive Leg” gesture was adjusted to remove its ability to track the active leg but keeping the ability to track movement in the passive leg. Pilot testing identified that the Kinect sensor was not able to accurately track the active leg movement during performance. The same occurred for the “Passive Arm” gesture and its ability to track the movement of the active arm during performance.

Therefore, the “Passive Leg” and “Passive Arm” gestures assessed only the passive limbs and checked for movement in either during subtest performance on the y-axis. The “Spine Angle” gesture was designed to check the angle of the spine between the hip and shoulder joints, allowing movement of up to 5° on the y-axis. This allowed prototype v4 to identify if a participant had excessive movement in the shoulder or hip region. The “Shoulder Angle” gesture was designed to check for excessive movement in the shoulders during test performance. However, during pilot testing, once the active shoulder was flexed, the shoulder joint moved and led to a fault being incorrectly identified. This gesture was removed as the range of movement at the joint varied considerably between participants. The “Hip Angle” gesture checked for movement between the left and right hips on the x-axis, with an acceptable range of 5° set. This ensured that any sway in the hips or torso region was detected by prototype v4. As per all other subtests, the tester had the option to manually check the “Pain” box on the user interface, which indicated pain was felt during performance and a score of zero was given by prototype v4. More detail regarding how prototype v4 distinguishes between different set up positions to accurately score performance can be found in Appendix H.

Corrective Exercises

As discussed earlier in this chapter, additional functionality was built in to prototype v4 in the form of an “Exercises” page. This allowed the software to automatically prioritise which weakness should be focussed on within a corrective exercise programme. This made prototype v4 more useful in applied settings and helped to ensure that the results of FMS performance were used in an effective manner. The FMS Manual provides clear guidelines for corrective exercise programming following screening, to ensure that testers know how to correctly interpret test results. Figure 6.6. provides details of the Corrective Strategy Algorithm included in the FMS Manual.

Corrective Strategy Algorithm

An algorithm, as defined by Cormen, Leiserson, Reivest and Stein in Introduction to Algorithms, is "any well-defined procedure describing how to carry out a particular task." Within the FMS there is an algorithm or procedure for addressing the "weak link" found using the screen.

Remember that you don't have to fix "everything" but rather the algorithm should point you towards the "one" thing you need to address as your priority.

While this may seem like a long process it takes less than a second or two to look at a score sheet and apply the algorithm.

For example, if the raw scores on a score sheet are as follows:
DS – 1
HS – 2/2
ILL – 1/2
SM – 2/2
ASLR – 1/1
TSPU – 1
RS – 2/2

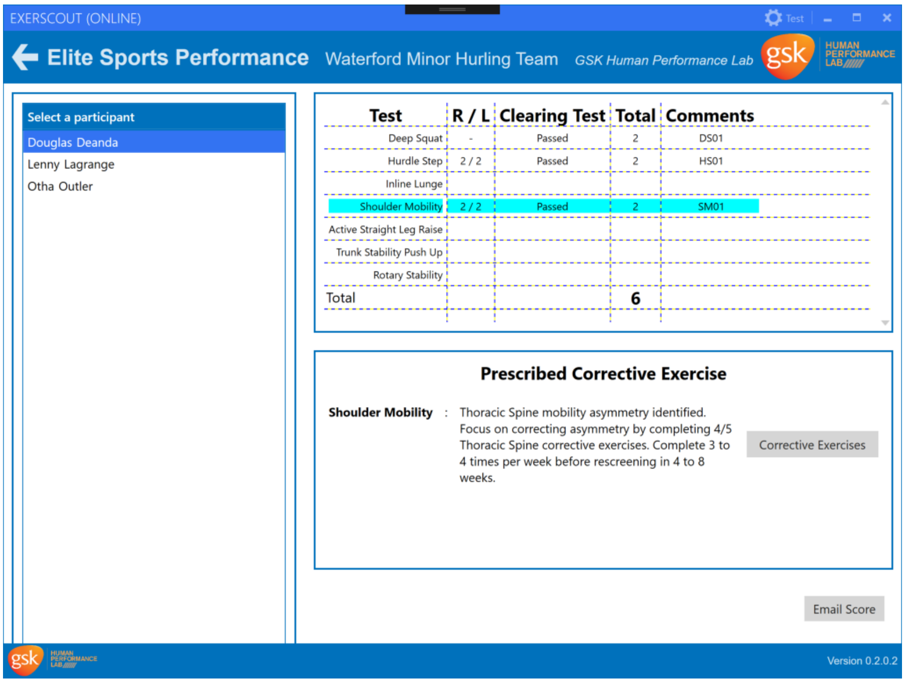
Following the order of the algorithm, look for scores of 1 or an asymmetry in order to identify the "weak link". (Remember the order of the score sheet is designed for efficiently administering the screen. The order of the algorithm is different based on the priorities for the corrective process.) So for this example, your eyes should immediately go to the ASLR and stop there because you have found your "weak link" in the order designated by the algorithm and don't need to look any further. Ignore all other scores and address the ASLR.

General Scoring Rules

1. A score of Zero must be referred to the appropriate medical professional.
    a. It is extremely important to establish a referral network of professionals – Physician, physical therapist, chiropractor, massage therapist, sports psychologist, etc. One thing to actively seek out in your immediate area is an individual trained in the SFMA where you can send those patients that score a 0 to be further evaluated based on a movement-based diagnostic system sharing the functional movement system philosophy.
2. Mobility patterns are addressed first because Stability/Motor Control cannot be present with reduced mobility. (Mobility must be restored before addressing stability or motor control. Appropriate levels of mobility ensure that adequate sensory input is being used to develop the appropriate stabilization strategies and the appropriate levels of motor control. Without quality levels of mobility, stability and motor control cannot and will not be maximized.)
3. A score of 21 is not the goal. The goal is to set a baseline and to work towards achieving at least 2's on each movement screen.

Figure 6.6. The Corrective Strategy Algorithm (Functional Movement Systems, 2017)

The FMS Corrective Strategy Algorithm shown in Figure 6.6 was used as the basis for the automated corrective exercise prescription process developed for prototype v4. Upon completion of the FMS, the scores for the last screening session completed by each participant were available within the software (Figure 6.7). This included each subtest score as well as the total FMS composite score.



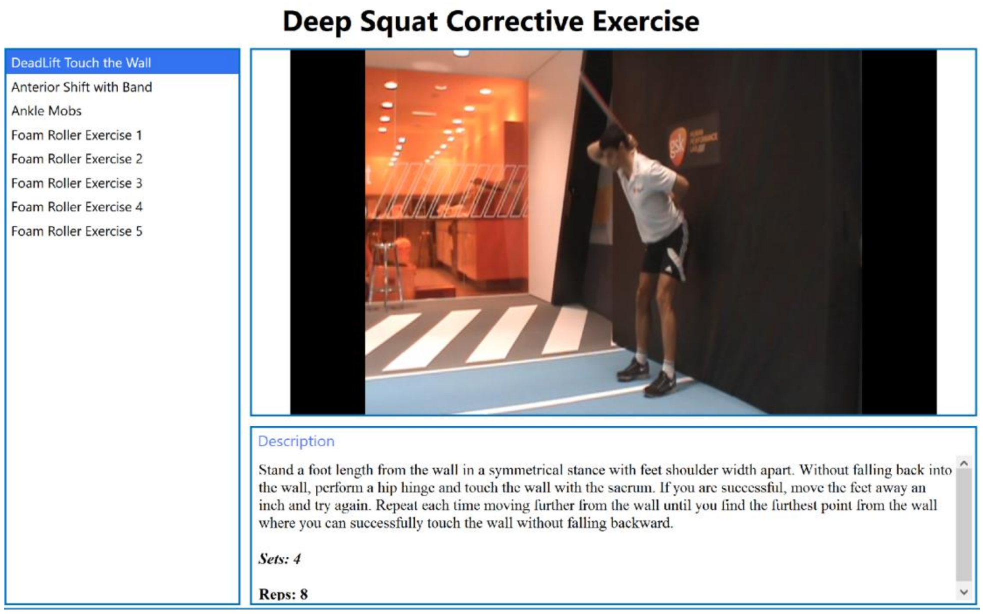
Test	R / L	Clearing Test	Total	Comments
Deep Squat	-	Passed	2	DS01
Hurdle Step	2 / 2	Passed	2	HS01
Inline Lunge				
Shoulder Mobility	2 / 2	Passed	2	SM01
Active Straight Leg Raise				
Trunk Stability Push Up				
Rotary Stability				
Total			6	

Prescribed Corrective Exercise

Shoulder Mobility : Thoracic Spine mobility asymmetry identified. Focus on correcting asymmetry by completing 4/5 Thoracic Spine corrective exercises. Complete 3 to 4 times per week before rescreening in 4 to 8 weeks.

Figure 6.7. The “Scores” page with corrective exercise recommendation for prototype v4

In addition to the scores, prototype v4 provided a breakdown of the key weaknesses to focus on in a corrective exercise intervention. Prototype v4 followed the FMS Corrective Strategy Algorithm to calculate which weakness to focus on. Any scores of zero were identified and it was recommended that the participant sought medical advice to fully understand what was causing this pain. If no pain was present for any subtests then any mobility issues were focussed on first, followed by stability weaknesses, and then finally any issues with total body patterns. The “Exercises” page within the user interface (Figure 6.8) provided details of the corrective exercise programme for each participant.


Deep Squat Corrective Exercise

Stand a foot length from the wall in a symmetrical stance with feet shoulder width apart. Without falling back into the wall, perform a hip hinge and touch the wall with the sacrum. If you are successful, move the feet away an inch and try again. Repeat each time moving further from the wall until you find the furthest point from the wall where you can successfully touch the wall without falling backward.

Sets: 4

Reps: 8

Figure 6.8. The “Exercise” page for prototype v4 showing a deep squat corrective exercise recommendation

The exercises included were selected based on the exercise library provided by the FMS website (Functional Movement Systems, 2017) and differed for each of the seven subtests to ensure that the weakness identified had the best opportunity to be strengthened. This automated approach provided users of the software with the ability to use the results generated by prototype v4 to provide a corrective exercise intervention to all participants, no matter their level of knowledge or expertise. It also allowed the intervention programme to be easily amended or updated after every screening session, to ensure the correct weaknesses were being worked on, and to check how effective the corrective exercise interventions were in improving an individual’s fundamental movement.

Software Operation

For the deep squat, hurdle step and inline lunge subtests, the user interface was operated in the same way for prototype v4 as that specified for prototype v3. A similar operational process was adopted for the remaining four subtests and is outlined in detail in Appendix H.

6.4 METHODS (PROTOTYPE V4 VALIDATION)

6.4.1 SAMPLE SIZE

Development of prototype v4 took 19 weeks in total. Upon completion of the development phase, a validation protocol was undertaken to assess the validity of the prototype v4 and its test-retest reliability. A sample of 16 healthy male and seven healthy female participants (mean±SD: age 22.83±5.33 years old; height 174.2±6.5 cm; body mass 72.23±11.6 kg) took part in the study. Ethical approval for the study was obtained from Waterford Institute of Technology Ethics committee, and each participant was asked to complete an approved Informed Consent form (Appendix A) and Pre-Test Questionnaire (Appendix B). Any participant who was unable to participate in their chosen sport at the time of screening due to injury or illness was excluded from participating in the study.

6.4.2 TESTING

Each participant was instructed to wear a T-shirt, shorts and footwear that they would normally wear to partake in exercise. Prototype v4 was set up on a Lenovo laptop running the Windows 10 operating system, which was linked to the Kinect sensor v2. The sensor was placed on a camera tripod and set at a height of 1.75 m. The FMS kit, which consisted of a 6ft by 2ft board, a 4ft dowel and rubber cord, was set up 4.00 m from the Kinect sensor v2 in the frontal plane. For the inline lunge subtest, the Kinect sensor was moved to a 30° angle from the FMS kit, but still 4.00 m from the sensor. Two Apple IPad's Version 2 (Apple Inc, 2015) were set up in the frontal and sagittal plane at a height of 1.00 m and distance of 4.00 m from the participant (Figure 6.9).

Figure 6.9. Set up of IPad’s and Kinect sensor for prototype v4 validation

A research assistant controlled prototype v4 during screening which automatically scored each participants performance on the seven subtests. Prior to each participant completing each subtest, a certified FMS tester instructed the participant in how to set up for each performance. Once set up was completed, the research assistant controlled subtest performance and instructed the participant when to start and finish each subtest performance. The certified FMS tester scored performance live using the recommended FMS scoresheet in the FMS Manual (Appendix I) and remained blinded to prototype v4 scores during screening.

Each participant completed each subtest three times, and any that involved unilateral assessment, three times on each side (right and left). A strict protocol was put in place by the researcher to ensure each subtest was conducted in a controlled manner and to ensure that the chance of bias was kept as low as possible. The protocols for the deep squat, hurdle step and inline lunge subtests remained as per the validation of prototype v3, described in chapter five.

To ensure that prototype v4 had every opportunity to provide the correct score during the deep squat, trunk stability and rotary stability subtests, the protocol provided them with only three attempts in total. The process for the deep squat remained as per prototype v3 validation. For the trunk stability subtest, each participant completed one attempt with their hands in the first set up position as per the FMS Manual, hence giving them the opportunity to score a three, one or zero, and then two attempts with their hands in the second set up position, giving them the opportunity to score a two, one or zero (6.10). For the rotary stability subtest each participant completed one attempt at scoring a three on each side by raising the arm and leg unilaterally, as per the recommended set up in the FMS Manual. Each participant was then given two opportunities to score a two by setting up for diagonal movement with the opposite arm and leg raised, as per the FMS Manual. This meant it was possible to be scored as a three, two, one or a zero on each side (Figure 6.11). Prototype v4 recorded all test performances so the researcher could take the best score from the three attempts following test completion. Although these processes did not strictly follow the FMS Manual (Functional Movement Systems, 2017) instructions, as the quality of performance was less relevant for the purposes of the test, it was felt that this was the most time efficient procedure that would reduce the chance of bias and provide prototype v4 with the opportunity to provide all four scoring outcomes for the three subtests. As the set up for the remaining four subtests do not change depending on performance, the standard FMS Manual protocols were followed for these subtests.

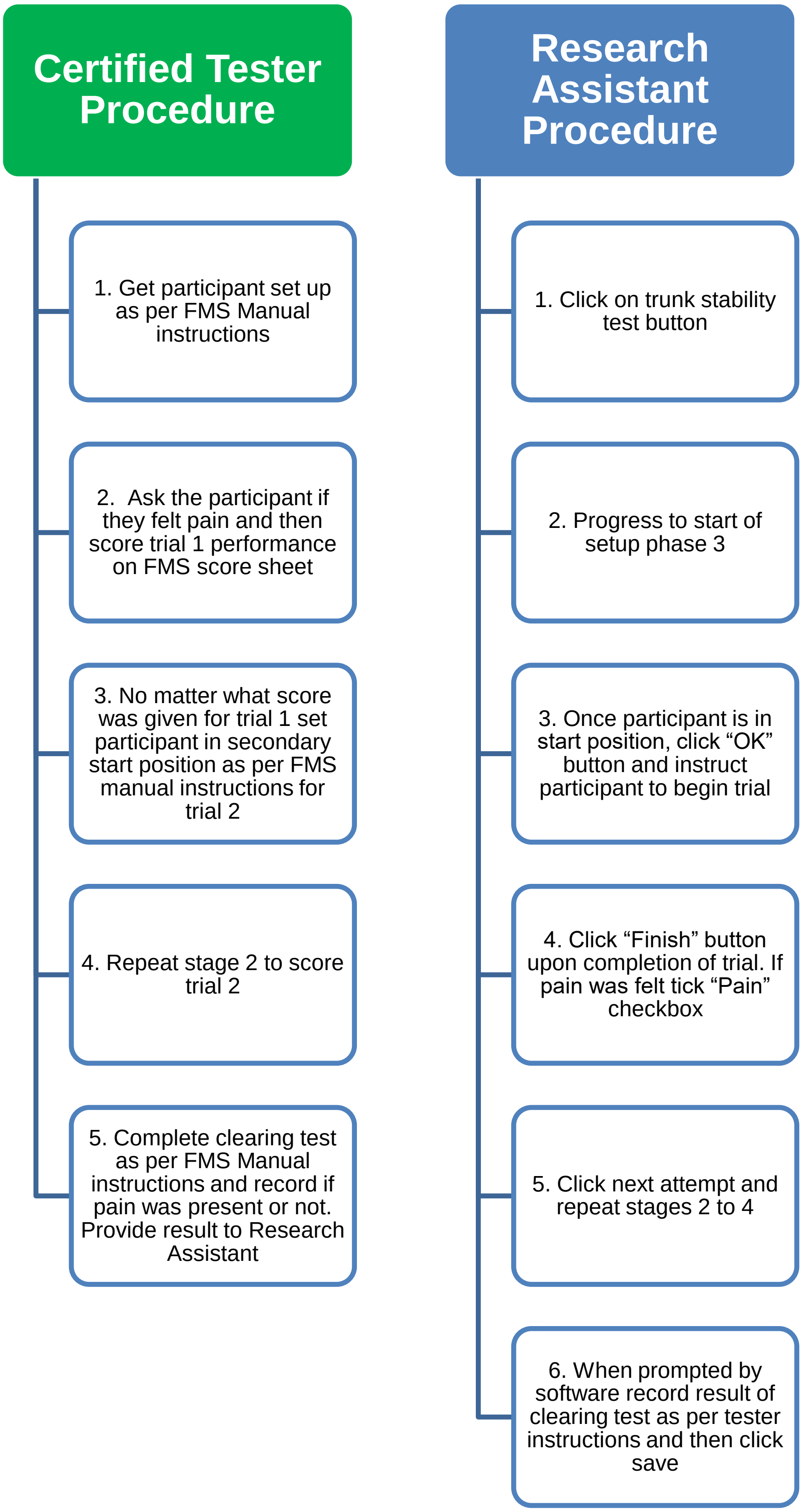


Figure 6.10. Protocol employed for validation of prototype v4 when assessing trunk stability performance

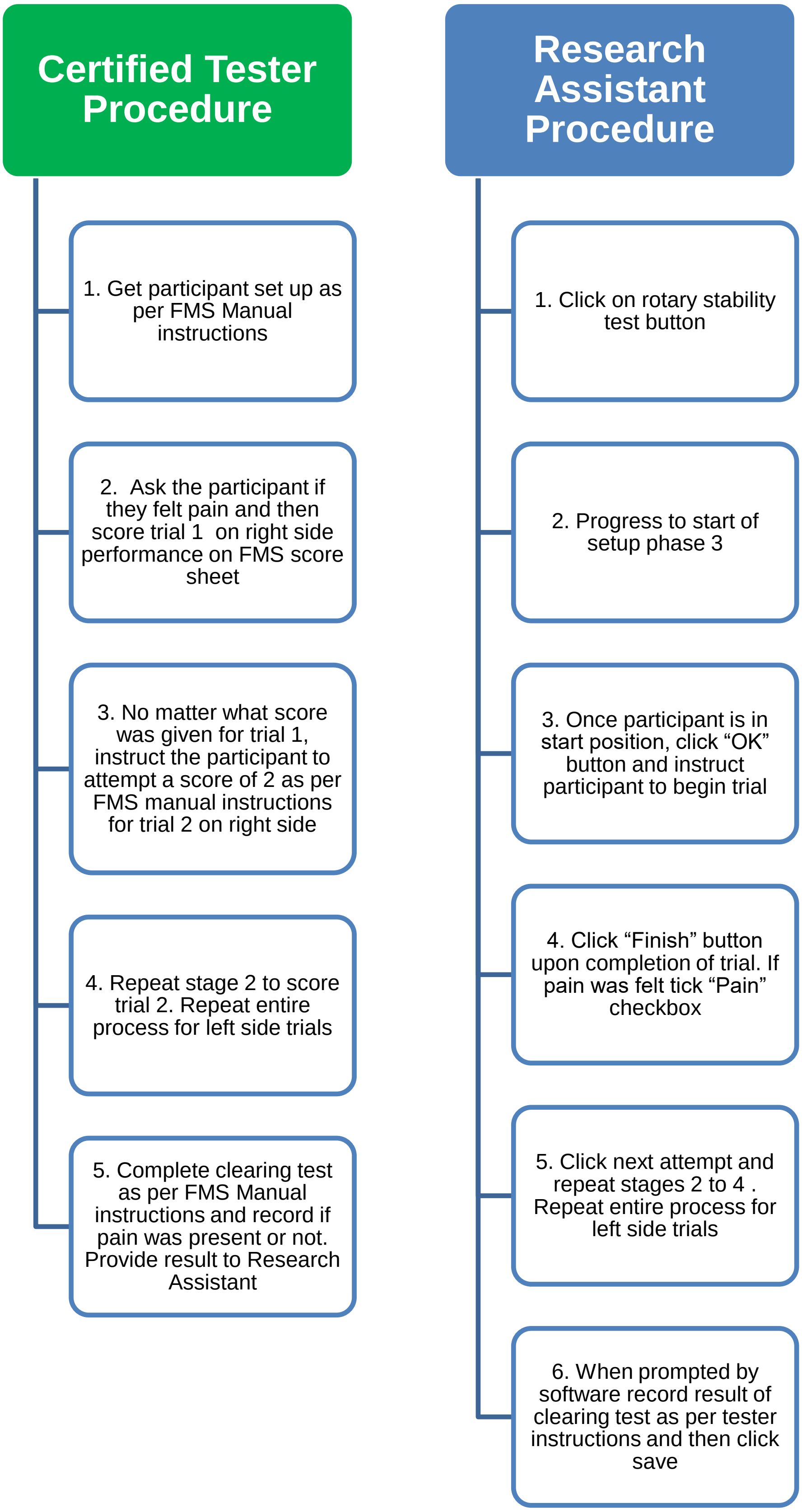


Figure 6.11. Protocol employed for validation of prototype v4 when assessing rotary stability performance

To assess test-retest reliability, nine of the participants, five males and four females (mean±SD: age 19.6±2.27 years old; height 177.7±4.9 cm; body mass 70.4±15.8 kg) repeated the seven FMS subtests again, one week after testing session one. The protocol followed was the same as that outlined in previous paragraphs. Due to time constraints placed on the participants, it was not possible to retest the remaining 14 participants a second time. The footage that was recorded by the Apple IPads was edited so that each participant's performance in both the frontal and sagittal planes across both testing sessions were converted into one video per performance (Figure 6.12). This footage was observed one week after the last testing session by a second certified FMS tester. The tester was asked to watch the footage in real time, without pausing or rewinding any performances, and to provide a score for each subtest performance on the Video Scoresheet (Appendix I). The tester was also asked to record the subtest that they would recommend being the focus of any corrective exercise programming for each participant on the same sheet.

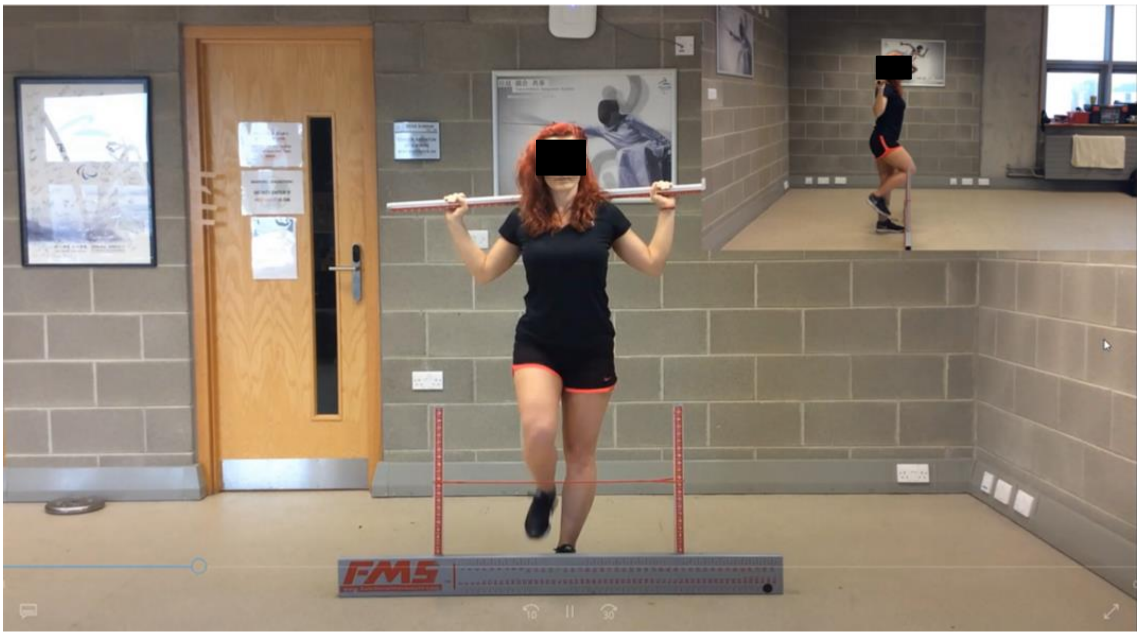


Figure 6.12. Example of video footage used by second certified FMS tester to score FMS performance for prototype v4 validation

Each participant was set up as a new user, under their participant code, within prototype v4 database. This meant that all scores automatically calculated by prototype v4 were stored following subtest performance, along with the subtest that the prototype automatically selected to be the focus of corrective exercise prescription. These scores were manually recorded on to the Master Scoresheet (Appendix I) by the researcher. The manual scores and corrective exercise recommendations recorded by the certified FMS tester who scored performance live, were manually added to the same sheet prior to the software scores being inserted. Finally, the scores and corrective exercise recommendations made by the certified FMS tester who scored the video footage were added to the same master document.

6.4.3 STATISTICAL ANALYSIS

To ascertain if prototype v4 could accurately assess the seven FMS subtests, statistical analysis was performed using IBM SPSS software version 24. Mean subtest scores and total FMS composite scores across all participants were calculated for prototype v4, manual live and manual video scoring methods for both testing sessions. To assess inter-rater reliability between certified FMS raters, a weighted Kappa (K_w) and agreement percentages between the two raters were calculated for all subtests, total FMS composite scores and corrective exercise recommendations. To assess the validity of prototype v4 as a method to assess all seven subtest performances and to recommend suitable corrective exercises, the absolute and percentage agreement between prototype v4 and both manual scoring methods for all participant's scores and corrective exercise recommendations were calculated. Additionally, a Spearman rank-order correlation test was conducted to understand the level of correlation between prototype v4 and both manual scoring methods, as another measure to test the validity of the software.

To establish the level of agreement between the scoring methods when calculating the total composite score for all seven subtests, a Bland Altman Limits of Agreement plot was constructed (Bland & Altman, 1986). To analyse test-retest reliability of the prototype v4 an Intraclass correlation coefficient (ICC) statistic was calculated for the seven subtest scores, total FMS composite score and the corrective exercises recommended by the prototype v4 for each testing session, using an alpha level of $p<0.05$. ICC values of 0.75 and above represent good reliability, those between 0.50 and 0.74 represent moderate reliability, and those below 0.50 indicate poor reliability (Teyhan et al., 2012).

6.5 RESULTS (PROTOTYPE V4)

Twenty-three participants were screened on all seven FMS subtests during session one. Mean subtest scores recorded by the three different scoring methods are outlined in Table 6.8. Mean total composite scores for all seven FMS subtests were 11.6 (±2.6) for manual live scoring, 11.4 (±1.7) for prototype v4 scoring and 11.8 (±2.2) for manual video scoring.

Table 6.8. Mean subtests scores recorded by the three different scoring methods during validation of prototype v4

	Deep Squat	Hurdle Step	Inline Lunge	Shoulder Mobility	Active Straight Leg Raise	Trunk Stability	Rotary Stability
Live Manual Scoring	1.5 (±0.6)	1.9 (±0.5)	1.9 (±0.6)	2.4 (±0.9)	1.4 (±0.6)	1.3 (±0.8)	1.2 (±0.4)
Prototype v4 Scoring	1.4 (±0.6)	1.9 (±0.3)	1.8 (±0.5)	2.4 (±0.9)	1.7 (±0.7)	1.0 (±0.2)	1.2 (±0.4)
Video Manual Scoring	1.5 (±0.6)	2.0 (±0.4)	1.8 (±0.6)	2.4 (±0.9)	1.6 (±0.7)	1.5 (±0.7)	1.3 (±0.5)

Live v Video manual scoring

The mean percentage agreement between the two different manual scoring methods was 94% across the seven subtests. The lowest percentage agreement was 91.3% for four subtests and the highest was 100% for the deep squat and shoulder mobility subtests (Table 6.9). When considering the weighted kappa levels, the mean value across all seven subtests was excellent (K_w=0.85). The lowest reported value was for the hurdle step test (K_w=0.71) which indicates good levels of agreement, and the highest reported value was for the deep squat and shoulder mobility tests (K_w=1.00) which indicates perfect agreement between the two raters.

Table 6.9. Level of agreement between live and video manual scoring for all seven FMS subtests during validation of prototype v4

	Agreement between live and video manual scoring		
FMS Subtests	**Number of participants**	**%**	**Weighted Kappa (K_w)**
Deep Squat	23	100	1
Hurdle Step	21	91.3	0.71
Inline Lunge	21	91.3	0.76
Shoulder Mobility	23	100	1
Active Straight Leg Raise	21	91.3	0.76
Trunk Stability	21	91.3	0.82
Rotary Stability	22	95.7	0.88
Mean for all subtests	21.7	94	0.85

N = number of participants out of 23 who received the same score from both scoring methods

When considering the total FMS composite scores, live manual scoring recorded a range of 7 to 18 for all participants compared to a range of 9 to 18 for video manual scoring (Figure 6.13). Mean agreement for the total FMS composite scores was 86.9% between live and video manual scoring with 20 out of 23 participants provided with the same score by both methods. Good levels of inter-rater reliability for total FMS composite score between the two scoring methods were recorded (ICC=0.90; 95% CI, 0.78 to 0.96; p<0.02).

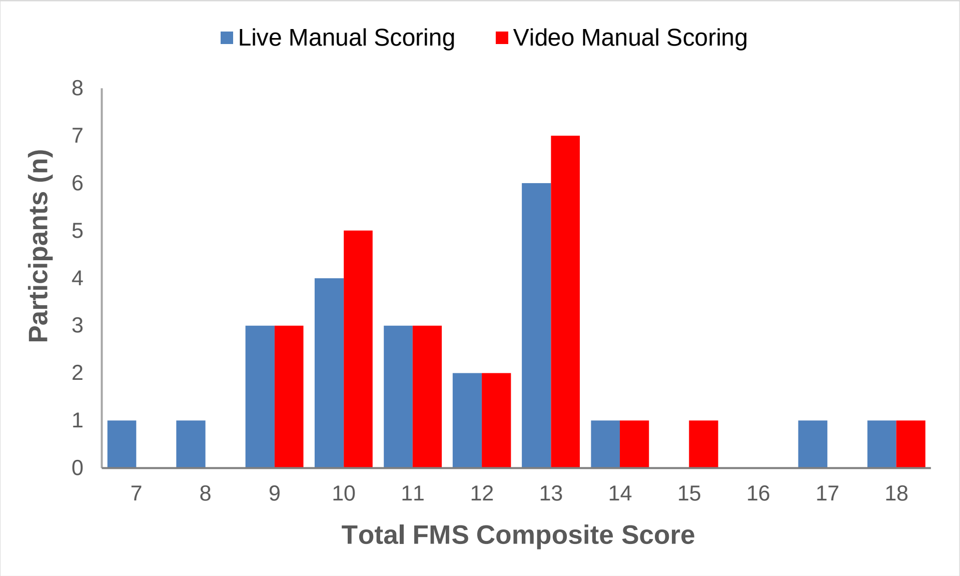


Figure 6.13. Absolute agreements between manual scoring for total FMS composite score

For corrective exercise recommendations, mean percentage agreement was good (86.9%) with 20 out of 23 participants provided with the same corrective exercise recommendation by the two scoring methods. A 100% agreement was recorded between the two scoring methods for four of the corrective exercise recommendations available (Table 6.10). The lowest agreement was seen for the shoulder mobility (50%) and rotary stability (50%) corrective exercise recommendations. Inter-rater reliability for corrective exercise recommendation between the two scoring methods was good (K_w=0.78).

Table 6.10. Corrective exercise recommendation agreements between live and video manual scoring methods

Corrective Exercise	Live Manual Scoring (n)	Video Manual Scoring (n)	Agreement (%)
Deep Squat	1	1	100
Hurdle Step	0	0	100
Inline Lunge	0	0	100
Shoulder Mobility	2	1	50
Active Straight Leg Raise	12	11	92
Trunk Stability	3	2	67
Rotary Stability	3	6	50
Medical Advice	2	2	100

Live manual scoring v Prototype v4 scoring

Mean agreement between live manual and prototype v4 scoring methods was 88% for all seven subtests. The active straight leg raise and trunk stability subtest reported the lowest agreement (78.3%), whilst the highest agreement was seen for the inline lunge and rotary stability subtests (95.7%) (Table 6.11 and 6.12). When considering correlation between the two scoring methods, a significant positive correlation was reported for all seven subtests, with the lowest correlation reported for the trunk stability subtest (r_s=0.46; $p<0.01$) and the highest correlation recorded for the shoulder mobility subtest (r_s=0.90; $p<0.01$).

Table 6.11. Level of agreement and correlation between live manual and prototype v4 scoring for all seven FMS subtests

	Agreement between manual scoring and prototype v4 scoring		
FMS Subtests	**N**	**%**	**Correlation (r_s)**
Deep Squat	20	86.9	0.79*
Hurdle Step	21	91.3	0.88*
Inline Lunge	22	95.7	0.54*
Shoulder Mobility	21	91.3	0.90*
Active Straight Leg Raise	18	78.3	0.73*
Trunk Stability	18	78.3	0.46*
Rotary Stability	22	95.7	0.87*
Mean for all subtests	20.3	88	0.74

*p<0.01, N = number of participants out of 23 correctly scored by prototype v4 compared to live manual scoring

Table 6.12. Actual agreement for all FMS scoring options between live manual and prototype v4 scoring for all seven FMS subtests (Figures in red show where agreement occurred between scoring methods)

Deep Squat					
	Live Manual Scoring				
Prototype v4 Scoring	**0**	**1**	**2**	**3**	**Total**
0	0	0	0	0	0
1	0	12	3	0	15
2	0	0	7	0	7
3	0	0	0	1	1
Total	0	12	10	1	23
Inline Lunge					
0	1	0	0	0	1
1	0	2	0	0	2
2	0	0	19	1	20
3	0	0	0	0	0
Total	1	2	19	1	23
Active Straight Leg Raise					
0	0	0	0	0	0
1	0	9	0	0	9
2	0	5	6	0	11
3	0	0	2	1	3
Total	0	14	8	1	23
Rotary Stability					
0	0	0	0	0	0
1	0	18	1	0	19
2	0	0	4	0	4
3	0	0	0	0	0
Total	0	18	5	0	23

Hurdle Step					
	Live Manual Scoring				
Prototype v4 Scoring	**0**	**1**	**2**	**3**	**Total**
0	0	0	0	0	0
1	0	3	0	0	3
2	0	1	18	1	20
3	0	0	0	0	0
Total	0	4	18	1	23
Shoulder Mobility					
0	1	0	0	0	1
1	0	3	0	0	3
2	0	0	4	0	4
3	0	0	2	13	15
Total	1	3	6	13	23
Trunk Stability					
0	1	0	0	0	1
1	0	17	2	3	22
2	0	0	0	0	0
3	0	0	0	0	0
Total	1	17	2	3	23
All Subtests					
0	3	0	0	0	3
1	0	64	6	0	70
2	0	6	58	2	66
3	0	0	4	15	19
Total	3	70	68	17	158

For total FMS composite scores prototype v4 scoring provided a range of scores from 8 to 15 compared to a range of 7 to 18 recorded by live manual scoring (Figure 6.14). Mean agreement for the total FMS composite scores was 78.2% between live manual and prototype v4 scoring with 18 out of 23 participants provided with the same score by both methods. Inter-rater reliability for total composite FMS score was good between the two scoring methods (ICC=0.82; 95% CI, 0.62 to 0.92; p<0.02).

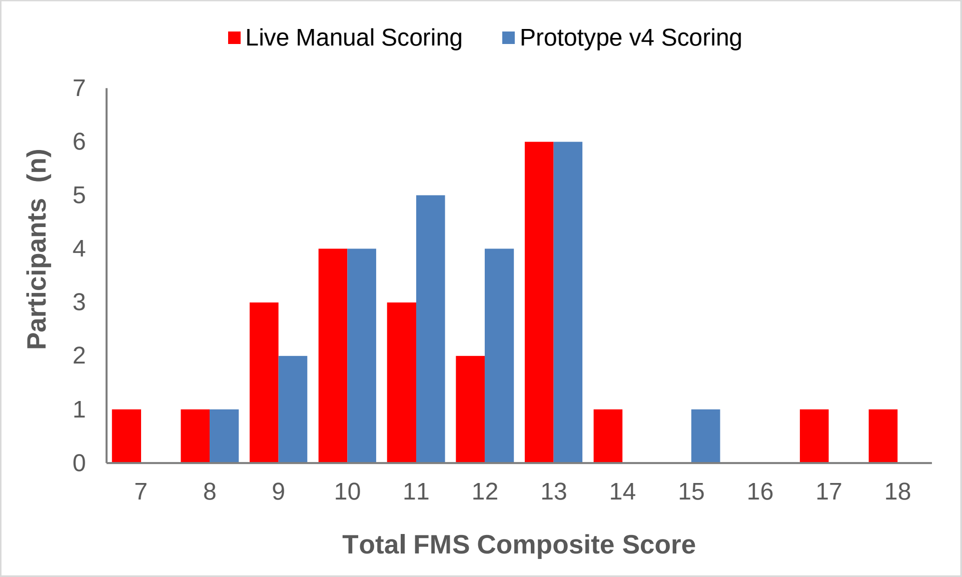


Figure 6.14. Absolute agreements between live manual and prototype v4 scoring methods for total FMS composite score

A Bland Altman plot was constructed (Figure 6.15) to assess the degree of difference between the scoring methods and the level of bias when recording total FMS composite scores. The difference between the total composite FMS scores identified by the two scoring methods was calculated for each participant. In addition, the mean total FMS composite score between the two scoring methods was established, for each participant. This provided a range of scores between 7.5 and 16.5 and resulted in 16 data points being visible on the Bland Altman Plot from the 23 scores used to calculate the limits of agreement. Using the standard deviation and the mean bias of the difference between scores of 1.33 and 0.17 respectively, 95% lower limits of agreement of -2.44 and upper limits of agreement of 2.79 were created.

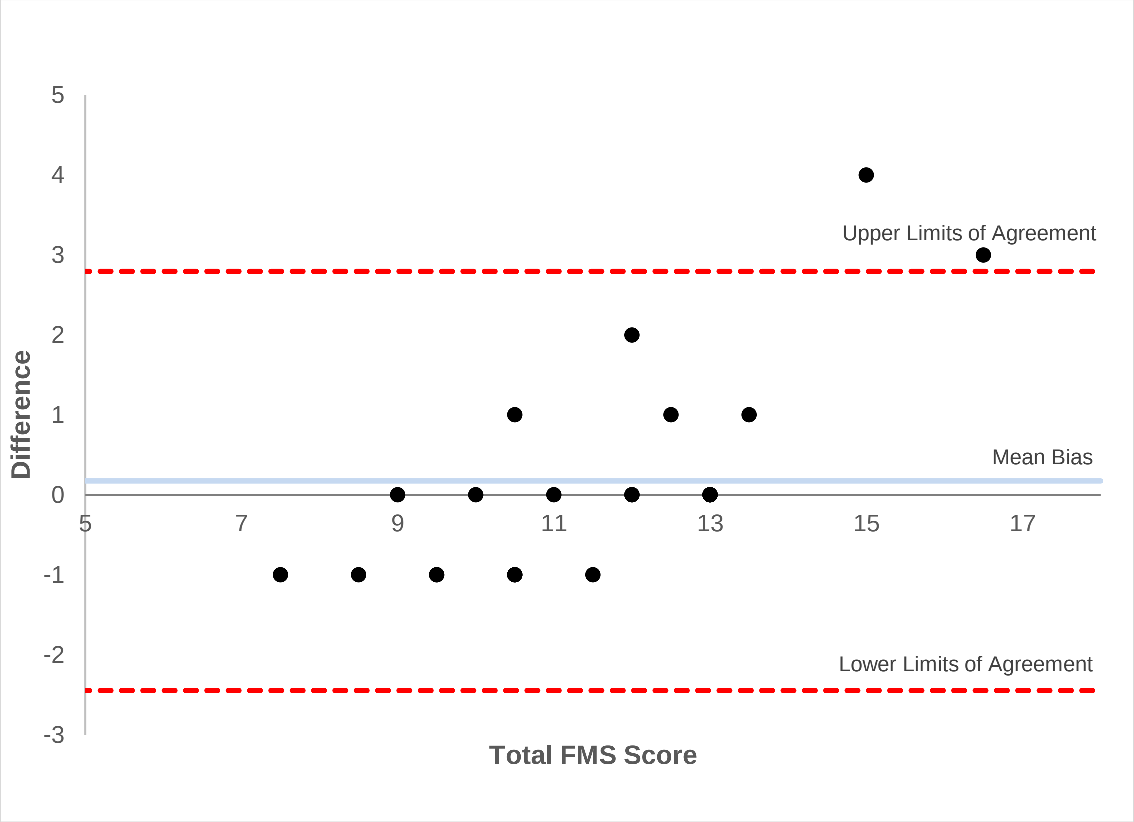


Figure 6.15. Bland Altman plot showing degree of difference between live manual and prototype v4 for total FMS composite scores (n=23)

For corrective exercise recommendations, the mean percentage agreement was good (78.2%) with 18 out of 23 participants provided with the same corrective exercise recommendation by the two scoring methods. A 100% agreement was recorded between the two scoring methods for three of the corrective exercise recommendations available (Table 6.13). No agreement was seen for the deep squat corrective exercise recommendation. Inter-rater reliability for corrective exercise recommendation between the two scoring methods was moderate (K_w=0.50).

Table 6.13. Corrective exercise recommendation agreements between live manual and prototype v4 scoring methods

Corrective Exercise	Live Manual Scoring (n)	Prototype v4 Scoring (n)	Agreement (%)
Deep Squat	1	0	0
Hurdle Step	0	0	100
Inline Lunge	0	0	100
Shoulder Mobility	2	2	100
Active Straight Leg Raise	12	7	59
Trunk Stability	3	5	60
Rotary Stability	3	7	43
Medical Advice	2	2	100

Video manual v Prototype v4 scoring

The mean percentage agreement between the two different scoring methods was 87% across the seven subtests. The lowest percentage agreement was 69.6% for the trunk stability subtest and the highest was 91.3% for four subtests (Table 6.14 and 6.15). When considering correlation between the two scoring methods, a significant positive correlation was reported for all seven subtests, with the lowest correlation reported for the trunk stability subtest (r_s=0.42; p=0.04) and the highest correlation recorded for the shoulder mobility subtest (r_s=0.90; p<0.01).

Table 6.14. Level of agreement and correlation between video manual and prototype v4 scoring for all seven FMS subtests

	Agreement between video and prototype v4 scoring		
FMS Subtests	**N**	**%**	**Correlation (r_s)**
Deep Squat	20	86.9	0.62*
Hurdle Step	21	91.3	0.86*
Inline Lunge	20	86.9	0.46**
Shoulder Mobility	21	91.3	0.90*
Active Straight Leg Raise	21	91.3	0.65*
Trunk Stability	16	69.6	0.42**
Rotary Stability	21	91.3	0.77*
Mean for all subtests	20	87	0.67

*p<0.01; **p<0.05, N = number of participants out of 23 correctly scored by prototype v4 compared to video scoring

Table 6.15. Actual agreement for all FMS scoring options between video manual and prototype v4 scoring for all seven FMS subtests (Figures in red show where agreement occurred between scoring methods)

Deep Squat

	Video Manual Scoring				
Prototype v4 Scoring	**0**	**1**	**2**	**3**	**Total**
0	0	0	0	0	0
1	0	12	3	0	15
2	0	0	7	0	7
3	0	0	0	1	1
Total	0	12	10	1	23

Inline Lunge

	Video Manual Scoring				
Prototype v4 Scoring	**0**	**1**	**2**	**3**	**Total**
0	1	0	0	0	1
1	0	2	0	0	2
2	0	2	17	1	20
3	0	0	0	0	0
Total	1	4	17	1	23

Active Straight Leg Raise

	Video Manual Scoring				
Prototype v4 Scoring	**0**	**1**	**2**	**3**	**Total**
0	0	0	0	0	0
1	0	9	0	0	9
2	0	3	8	0	11
3	0	0	1	2	3
Total	0	12	9	2	23

Rotary Stability

	Video Manual Scoring				
Prototype v4 Scoring	**0**	**1**	**2**	**3**	**Total**
0	0	0	0	0	0
1	0	17	2	0	19
2	0	0	4	0	4
3	0	0	0	0	0
Total	0	17	6	0	23

Hurdle Step

	Video Manual Scoring				
Prototype v4 Scoring	**0**	**1**	**2**	**3**	**Total**
0	0	0	0	0	0
1	0	2	1	0	3
2	0	0	19	1	20
3	0	0	0	0	0
Total	0	2	20	1	23

Shoulder Mobility

	Video Manual Scoring				
Prototype v4 Scoring	**0**	**1**	**2**	**3**	**Total**
0	1	0	0	0	1
1	0	3	0	0	3
2	0	0	4	0	4
3	0	0	2	13	15
Total	1	3	6	13	23

Trunk Stability

	Video Manual Scoring				
Prototype v4 Scoring	**0**	**1**	**2**	**3**	**Total**
0	1	0	0	0	1
1	0	15	4	3	22
2	0	0	0	0	0
3	0	0	0	0	0
Total	1	15	4	3	23

All Subtests

	Video Manual Scoring				
Prototype v4 Scoring	**0**	**1**	**2**	**3**	**Total**
0	3	0	0	0	3
1	0	60	10	3	70
2	0	5	59	2	66
3	0	0	3	16	19
Total	3	65	72	21	161

For total FMS composite scores, prototype v4 scoring provided a range of scores from 8 to 15 compared to a range of 9 to 18 recorded by video manual scoring (Figure 6.16). Mean agreement for the total FMS composite scores was 78.2% between video manual and prototype v4 scoring with 18 out of 23 participants provided with the same score by both methods. Inter-rater reliability for total composite FMS score was good between the two scoring methods (ICC=0.77; 95% CI, 0.54 to 0.90; p<0.01).

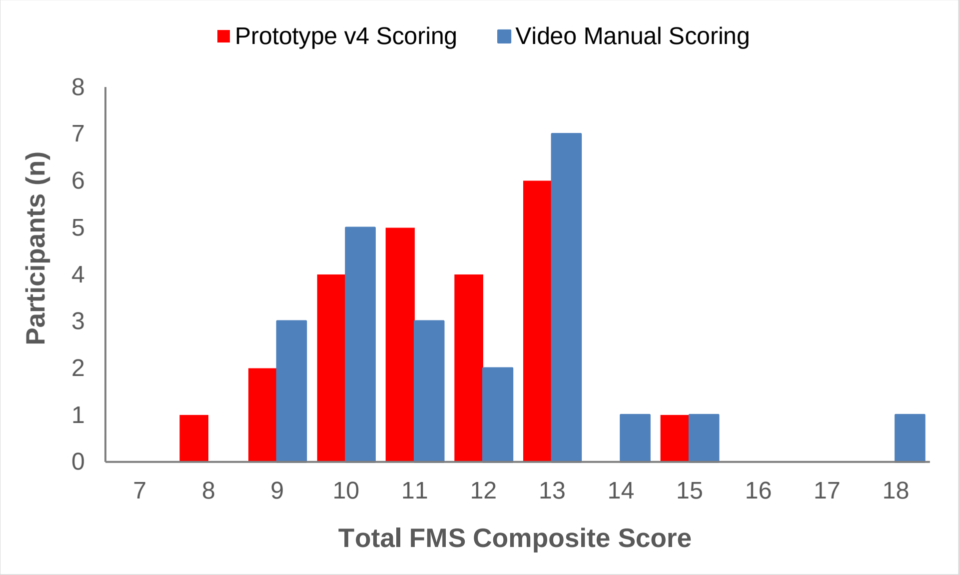


Figure 6.16. Absolute agreements between prototype v4 and video manual scoring methods for total FMS composite score

A Bland Altman plot was constructed (Figure 6.17) to calculate the difference between the total composite FMS scores identified by the two scoring methods for each participant. In addition, the mean total FMS composite score between the two scoring methods was established, for each participant. This provided a range of scores between 9 and 16.5 and resulted in 16 data points being visible on the Bland Altman Plot from the 23 scores used to calculate the limits of agreement. Using the standard deviation and the mean bias of the difference between scores of 1.27 and 0.43 respectively, 95% lower limits of agreement of -2.06 and upper limits of agreement of 2.93 were created.

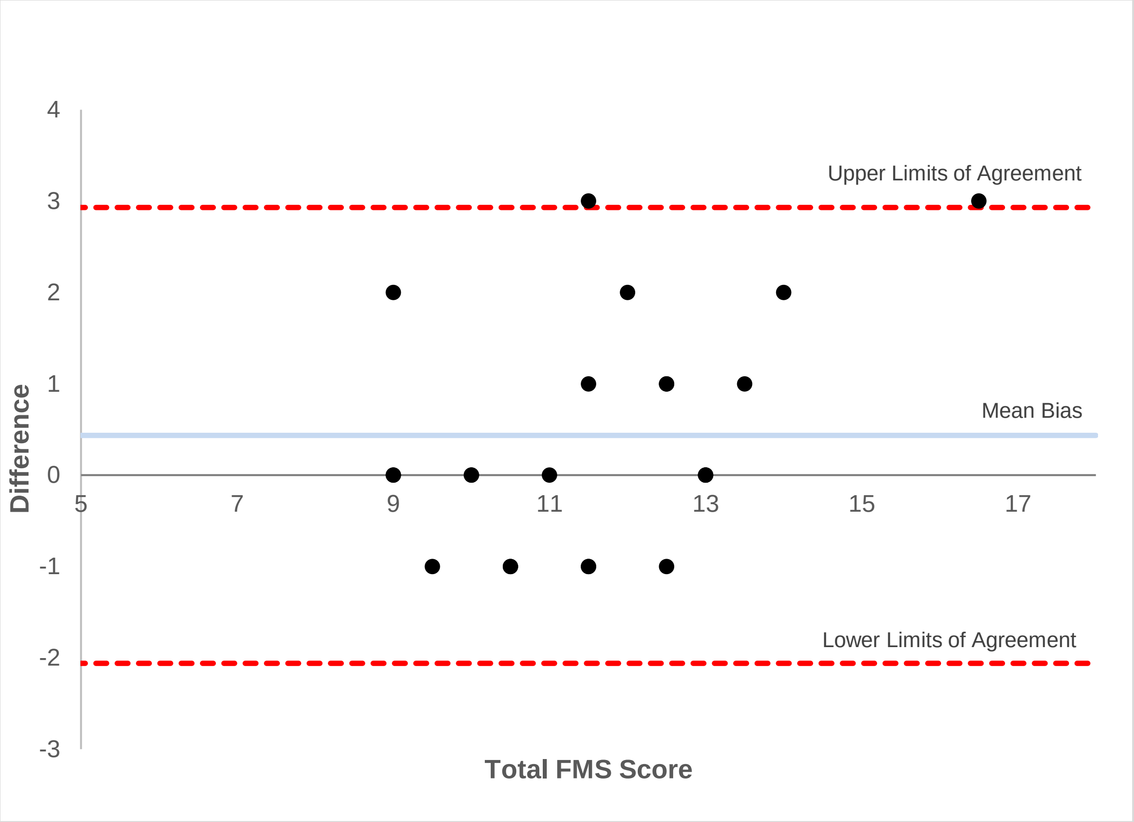


Figure 6.17. Bland Altman plot showing degree of difference between video manual and prototype v4 total FMS composite scores (n=23)

For corrective exercise recommendations, the mean percentage agreement was moderate (78.2%) with 18 out of 23 participants provided with the same corrective exercise recommendation by the two scoring methods. A 100% agreement was recorded between the two scoring methods for three of the corrective exercise recommendations available (Table 6.16). No agreement was seen for the deep squat corrective exercise recommendation. Inter-rater reliability for corrective exercise recommendation between the two scoring methods was moderate (K_w=0.55).

Table 6.16. Corrective exercise recommendation agreements between prototype v4 and video manual scoring methods

Corrective Exercise	Prototype v4 Scoring (n)	Video Manual Scoring (n)	Agreement (%)
Deep Squat	0	1	0
Hurdle Step	0	0	100
Inline Lunge	0	0	100
Shoulder Mobility	2	1	50
Active Straight Leg Raise	7	11	64
Trunk Stability	5	2	40
Rotary Stability	7	6	86
Medical Advice	2	2	100

Test-retest reliability

Nine of the participants were retested on all seven FMS subtests one week after the first testing session to assess the test-retest reliability of prototype v4. Mean test-retest reliability was excellent across all seven subtests (ICC=0.96). Five of the FMS subtests reported perfect test-retest reliability (ICC=1.00) between testing session one and two (Figure 6.18). The hurdle step subtest had the lowest test-retest reliability (ICC=0.78; 95% CI, 0.09 to 0.95; p=0.02). For total FMS composite score, test-retest reliability was good (ICC=0.97; 95% CI, 0.87 to 0.99; p<0.01) and was perfect for corrective exercise recommendations (ICC=1.00).

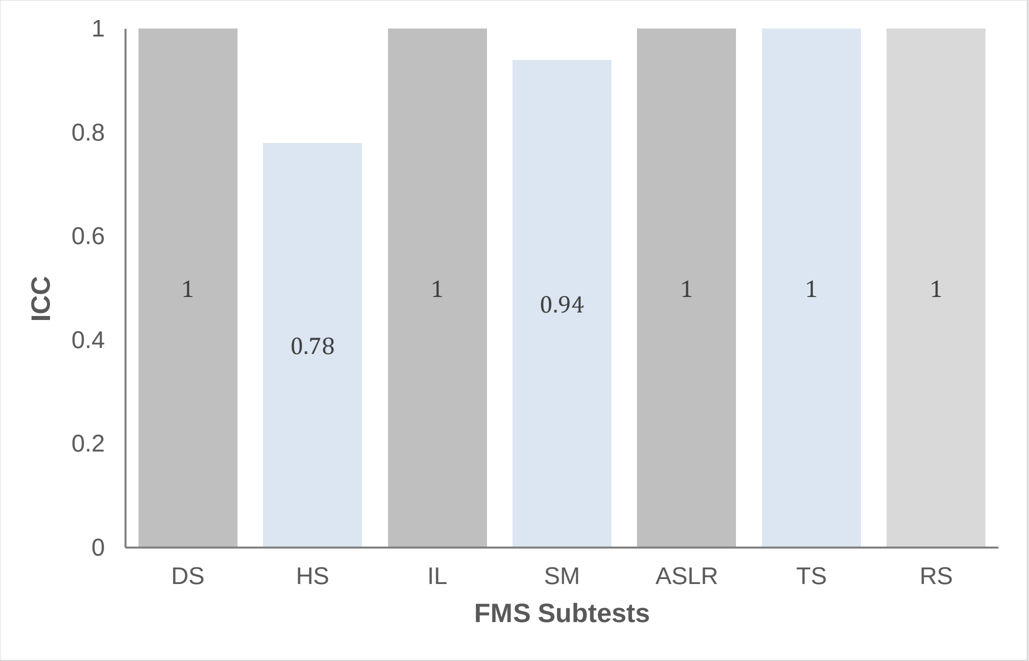


Figure 6.18. Test-retest reliability between testing session one and two for prototype v4

6.6 DISCUSSION (PROTOTYPE V4)

There were good levels of inter-rater reliability and levels of agreement when comparing live manual versus video manual scoring methods. The reported levels of reliability and agreement were similar for subtest scoring to that seen in earlier research comparing manual FMS raters (Parenteau et al., 2014; Schneiders et al., 2009). This suggests that both scoring manual methods could validly be used as the gold standard against which to compare prototype v4 automated scoring to assess its level of validity and reliability.

6.6.1 SCORING SYSTEM

Prototype v4 showed good levels of agreement and correlation across the seven subtests with both live and video manual scoring methods, suggesting it is a valid tool to use to assess certain FMS subtests compared to manual assessment. In particular, the hurdle step, shoulder mobility and rotary stability subtests reported higher levels of agreement between prototype v4 and both scoring methods to that reported in previous research in this area (Minick et al., 2010; Teyhen et al., 2012). Although these results are promising, it is still not possible to determine if the software would reduce the subjectivity of FMS scoring in applied settings. The manual FMS scoring system is subjective by nature and it therefore has limitations when used as a reference against which to compare the automated scoring method of prototype v4. Previous research by Whiteside et al. (2014) found similar difficulties when comparing manual FMS assessment to measurement of performance using inertial measurement units. These authors suggested that manual grading may not be a valid measurement instrument, based on low levels of agreement between the scoring methods when assessing six FMS subtests, which were all lower than the levels of agreement reported for prototype v4. Although these authors raise a valid argument regarding manual assessment, the FMS was designed to be manually assessed, and to date no suitable alternative objective method exists to score the seven subtests. Therefore, the results from this current study provide novel insight regarding how an automated system can be designed to attempt to reduce the subjectivity of the manual scoring system.

The FMS scoring criteria does not provide detailed kinematic information relating to joint positioning or angles when scoring subtests and therefore leaves a level of ambiguity when grading performance on all seven subtests. Each tester will apply their own interpretation of the criteria to each performance. In this study, the thresholds set for the custom gestures for prototype v4 were based on the researchers own interpretation of the criteria. Although the researcher is an experienced level 2 FMS tester, this approach could have positively impacted the levels of agreement between the live manual and prototype v4 scoring methods. Including a second tester in this validation provided a more objective measure against which to compare prototype v4 scoring, and agreement between these methods was good to excellent. The second tester was a very experienced level 2 tester, who teaches on FMS certification courses, and therefore their interpretation of performance should be closely matched with the FMS Manual scoring criteria. However, the assessment was still subjective and further highlights the difficulty faced in this programme of research in assessing the validity and reliability of prototype v4 against manual scoring.

6.6.2 THRESHOLDS

Overall the level of agreement and correlation between the scoring methods was moderate to excellent for all seven subtests. However, three of the subtests showed lower levels of agreement against both live and video manual scoring, suggesting additional work is required before prototype v4 could be classified as a valid method to score the FMS. The trunk stability subtest showed the lowest levels of agreement and correlation with both manual scoring methods. When analysing this in more detail it was apparent prototype v4 was unable to identify performances scoring a two or three, suggesting that the custom gestures set for this subtest were overly sensitive. It also indicates that the Kinect sensor itself had difficulty tracking joint centres during performance due to the prone position adopted by each participant. Whole body movement is assessed during performance and it was a challenge to ensure that the sensor could accurately identify the side of the body furthest from the sensor when performing the subtest. As a result, the Kinect sensors “Joint Orientation” system has great difficulty tracking certain joints of the body during test performance. This resulted in prototype v4 downgrading performance to a one on most occasions for the trunk stability subtest. Consideration needs to be given to the best position in which to set the Kinect sensor up for this subtest to maximise the opportunity to accurately track all body joints.

The inline lunge subtest showed good levels of agreement with both manual scoring methods, but low levels of correlation. Prototype v4 was unable to identify the one participant that was scored as a three by manual scoring methods, but it is difficult to assess whether prototype v4 could validly be used to identify performances scoring a three, due to the low number of threes scored in this sample. This was also the case during validation of prototype v3, where the majority of performances were scored as a two, and highlighted that further validation using a sample with a wider range of subtest scores would help to establish prototypes v4's accuracy levels when scoring performances of three on the inline lunge. Similarly to the trunk stability subtest, the inline lunge subtest required one side of the body to be facing away from the sensor, despite the fact the FMS board was placed at a 30° angle to the sensor. This made it more difficult for the Kinect to track the joints farthest from the sensor, which could have negatively impacted prototype v4's scoring accuracy. Further work assessing the most effective set up for the inline lunge could be worthwhile in an attempt to improve the accuracy of scoring performances of three.

For the deep squat subtest, despite levels of agreement and correlation being comparable to the other subtests, prototype v4 was unable to accurately score performances of two. Although it correctly identified the only participant who scored a three, it incorrectly scored three participants as a one, each of whom were scored as a two by both manual scoring methods. Throughout the validation process of the four prototypes, there were a high number of performances scored as a one by manual scoring methods. This resulted in good levels of agreement with prototype scoring, as a one is the easiest score to identify, and perhaps masked the prototypes weakness in scoring higher quality performances. Previous research by Jensen, Weilbrenner, Rott, & Eskofier (2013) assessed agreement between a semi-automated wireless inertial measurement system and manual assessment of the deep squat subtest in 10 subjects. Overall agreement was 80%, with two participants scored incorrectly by the system. However, the system correctly identified six out of seven performances scored as a two, suggesting it was a more valid method than prototype v4 at scoring mid-range performance on this subtest.

The limitations outlined in the previous paragraphs suggest that prototype v4 requires further work before it could be used in applied settings. Understanding the most suitable thresholds to apply to each subtest is an area that still needs further work. As the FMS scoring criteria is so simplistic, any study that attempts to develop a new measurement unit to automatically assess the FMS, must rely on the researcher taking the scoring criteria and developing thresholds, based on their interpretation of the criteria, to apply to their system. For prototype v4, the deep squat subtest included seven custom gestures that were based on the angle and alignment of key joints or body parts outlined in the FMS scoring criteria to assess deep squat performance. Previous research has used alternative thresholds, based on their own interpretation of the FMS scoring criteria, to assess deep squat performance. Whiteside et al. (2014) incorporated six criteria in their system, all of which focussed on the angle of different joints involved in deep squat performance, whilst Jensen et al. (2013) incorporated only four criteria into their semi-automated system, again using joint angle as the key measurement. These studies highlight how difficult it is for the FMS scoring criteria to be interpreted in a way that can easily be transferred to an automated system to provide accurate scoring assessment.

6.6.3 CORRECTIVE EXERCISES

Prototype v4 correctly assigned the corrective exercise programme it had been programmed to provide for all 23 participants based on the scores it gave for all subtests, and agreement with manual scoring methods, in regard to corrective exercises prescribed, was good for both live and video manual scoring. These results highlight that prototype v4 could confidently be used in applied settings by practitioners to prescribe corrective exercise interventions based on each individual’s greatest area of weakness in terms of FMS performance, if the validity of automated subtest scoring can be proven. Any differences in agreement between prototype v4 and both scoring methods regarding corrective exercises prescribed, were caused due to inaccurate subtest scoring by prototype v4. For example, many participants scored one on the trunk stability subtest, leading to several participants being prescribed exercises to improve weaknesses on this subtest by prototype v4. The FMS “Corrective Exercise Algorithm” directs testers to focus on stability issues if there are no scores of one on the mobility subtests. As such, the low levels of agreement on the trunk stability subtest, between prototype v4 and manual scoring methods, significantly affected corrective exercise programme prescription and reduced agreement between the scoring methods.

6.6.4 TEST-RETEST RELIABILITY

The test-retest reliability assessed prototype v4’s ability to reliably score the same participants on the FMS at two different timepoints, seven days apart. The results show that test-retest reliability was good, with a mean ICC of 0.95 and was higher than that previously reported for manual raters (Shultz et al., 2013; Smith et al., 2013). Five subtests had perfect test-retest reliability (ICC-1.00) with the shoulder mobility subtest (ICC=0.94) and the hurdle step (ICC=0.78) showing good test-retest reliability. The same high levels of test-retest reliability were also reported for total FMS composite score (ICC=0.97). Whilst these test-retest reliability results show great promise and highlight that prototype v4 is very consistent in its automated scoring methods, the size of the sample (n=9) is small and means the results must be viewed with some caution. Additional studies with larger sample sizes and screening taking place over multiple time points would provide a more comprehensive assessment of prototype v4's test-retest reliability. It would also be useful to compare test-retest reliability of manual scoring against prototype v4 scoring, to understand the level of reliability across different timepoints on the same sample of participants. Due to time constraints, this study was unable to assess manual test-retest reliability between the two different testing sessions and as such could not make such a comparison.

6.7 CONCLUSION

This research study has discussed the development of prototype v4 and attempted to validate its ability to automatically score all seven FMS subtests compared to live and video manual scoring. Small adjustments were made to the custom gestures for the deep squat, hurdle step and inline lunge subtests to improve prototypes v4's accuracy in scoring these subtests. In addition, the functionality of prototype v4 was enhanced to allow it to assess the remaining four subtests and to automatically prescribe corrective exercise programming based on the results of the FMS. The FMS Manual was used to provide the scoring criteria against which the custom gestures were developed for each of the four new subtests and formed the basis for the corrective exercise algorithm developed for the software. Prototype v4 was designed with the Kinect sensor set at a height of 1.75 m to maximise the sensor's chances of tracking joint centres during subtests that required the participant to be in a prone, supine or quadruped position.

Although the results suggest that prototype v4 is a valid method to assess the hurdle step, shoulder mobility, active straight leg raise, and rotary stability subtests compared to both live and video manual scoring, the results from the remaining subtests and total FMS composite scores, mean prototype v4 could not be used in applied settings to assess the FMS. For a new measurement method to be accepted by practitioners it must achieve a level of validity and reliability that exceeds existing assessment methods in all areas. Prototype v4 shows great promise as an FMS assessment tool and provides some novel and effective processes that help improve the efficiency of the FMS. However, it cannot provide the level of validity required when scoring every subtest and can therefore not be seen as a reliable alternative to manual rating. Further research in this area is worthy of consideration, with the aim of improving the accuracy of scoring on the deep squat, inline lunge and trunk stability subtests in particular. This includes comparing prototype v4, to alternative, objective, measurement units, to fully understand its accuracy in tracking joint movement throughout subtest performances, to provide a more robust measure of its overall validity and reliability when scoring the FMS. Additional work to develop standard thresholds to score each subtest, that could be adopted by both manual raters and alternative measurement units, would also provide useful insight into how to adapt existing prototype v4 custom gestures to improve validity and reliability. Finally, further work should also consider the most appropriate set up of the Kinect sensor for each subtest and whether the height and positioning should be adjusted to suit the individual requirements of each subtest, in an attempt to improve its scoring accuracy.

Including two testing sessions allowed the consistency of prototype v4’s scoring to be assessed to ensure it was able to provide reliable scoring across different testing sessions. The test-retest reliability of prototype v4 was very high and showed that it is very consistent when assessing the FMS and recommending corrective exercise programmes. This was very promising from an applied perspective, as practitioners need to be confident that the software provides the same level of accuracy for every single testing session. This study suggests that if the validity of prototype v4 could be improved for all seven subtests, the test-retest reliability is at a higher level than manual assessment and could therefore be a very effective tool in applied settings. Additional research with larger sample sizes and more testing sessions would be worthwhile to provide further clarification of prototype v4’s test-retest reliability.

6.8 REFLECTIONS

Chapters four to six outlined the development and validation process completed as part of this programme of research in an attempt to create automated software to assess the FMS. This process ultimately resulted in an automated system being created that had good levels of validity and reliability compared to manual scoring, but still had areas of weakness that would require future development work in order to improve the accuracy and consistency when scoring all seven FMS subtests. These weaknesses can partly be attributed to some of the challenges that were faced during development of the automated system, which in the main were a result of trying to replicate the FMS manual scoring system when creating the automated approach.

6.8.1 DEVELOPMENT CHALLENGES

Results from this programme of research highlighted that the custom gestures set for certain subtests required further work before the novel software could be seen as a reliable tool to use in applied settings. These custom gestures were set for different scores on each subtest and were based on the researchers own interpretation of the FMS scoring criteria. As the criteria is designed to be simplistic in its nature, setting detailed thresholds for each score on all seven subtests provided a significant challenge. These thresholds may well have been set differently if other expert FMS raters were building a new measurement device to assess the FMS, due to their interpretation of each scoring criteria, resulting in different measurements being used by the software to score each performance. This is highlighted in previous research that has attempted to develop automated devices to assess the FMS. Whiteside et al. (2014) incorporated six different criteria into their system for scoring the deep squat subtest, compared to just four criteria used by Jensen et al. (2013) in their semi- automated system to assess the same subtest. Therefore, further research should be considered, whereby the manual FMS assessments of several expert raters are observed and scrutinised to understand how each apply the FMS scoring criteria. Based on these observations, and input from the raters themselves, a set of detailed thresholds could be developed for the purposes of assessment by automated, objective, systems. Such standardised thresholds could be incorporated into the novel software used in this programme of research, in an attempt to improve the accuracy and reliability of scoring all seven FMS subtests. The FMS was designed to be a low cost, efficient and simple screening tool to assess individual’s movement patterns, used to guide the next phase of intervention or assessment. Therefore, any detailed thresholds that are developed should be used for the purposes of developing novel measurement systems, and not for the purposes of manual scoring of the FMS, to offer an objective, accurate and reliable alternative to manual FMS assessment.

In addition, certain FMS subtests themselves, did not lend themselves well to being assessed using depth camera technology such as the Kinect sensor. The Active Straight Leg Raise, Trunk Stability and Rotary Stability subtests involve a participant lying in a prone or supine position. This immediately causes difficulties for the Kinect sensor as it is unable to accurately track the joints of the body that are furthest from the sensor and has difficulty distinguishing between limbs on different side of the body. Therefore, a considerable amount of time was spent working on a solution to this issue to ensure the software was able to score these subtests in an accurate manner. The challenge was overcome through adopting a structured set up process for each of the subtests, to ensure the Kinect sensor was able to track the joints of the body nearest the sensor and was not attempting to track the joints furthest from the sensor. Whilst this allowed the automated approach to score the subtests it resulted in a more time consuming set up process for each of the three subtests, which from an applied perspective would not be as efficient as using manual scoring. Any future research that considers developing an automated approach to assessing the FMS, should consider utilising multiple depth cameras in order to improve the joint identification and tracking capabilities of the software, whilst ensuring the set up processes for each subtest do not change from those used during manual scoring to ensure they remain time efficient. A similar issue was found for the shoulder mobility subtest, whereby the sensor was unable to track the hands as soon as they were placed behind the body. This was overcome by adopting a semi-automated approach to this subtest, which although effective, was not desirable from the researcher’s perspective as it did not meet the initial objective of fully automating the FMS scoring process. A system that includes multiple cameras should be able to overcome this particular issue, as the different perspectives should ensure the hands can be tracked at all times.

6.8.2 VALIDATION LIMITATIONS

Although the challenges faced during development were overcome through a continuous process of testing and reiteration, the programme of research also had some limitations from a validation perspective, which had implications for the validity and reliability results.

The major limitation during all four validation studies was the limited spread of scores across the samples used. In particular, for the seven subtests there were a limited number of scores of threes reported, and in some cases, a particularly high number of ones recorded. This could have unduly affected the validity and reliability results from each of the studies as a performance that scores a one on each of the subtests is the easiest for a manual rater and the automated software to identify, thus increasing the chances of agreement between the two methods. For example, on the Hurdle Step subtest, a one is scored if the individual makes contact with the hurdle or loses their balance, both clearly identifiable actions that both systems would be able to detect with relative ease. Therefore, a sample with a higher proportion of performances scored as a one would likely have increased the agreement levels between the two methods compared to a sample with a more even spread of scores across all subtests. To achieve such a spread of scores is problematic, particularly when considering scores of threes on subtests such as the rotary stability and deep squat, which are difficult to achieve. However, future research in this area should consider either manipulating the performances of the sample in attempt to get an even spread of scores across the sample, or simply only analysing the same number of scores of one, two and three across a sample to ensure an even spread is being analysed.

A further limitation of the validation studies within the programme of research was the number and type of raters used to assess performance. Although the final validation included comparisons against two manual testers, further validation involving additional raters would be worthwhile. This should include both live manual raters as well as video based manual raters, with various levels of experience. Such validation would provide a clearer understanding of exactly how reliable and accurate the software is when scoring all seven FMS subtests and allow practitioners to evaluate whether the software is able to be used in applied settings. The validation completed in this programme of research used experienced FMS raters (both level 2 certified), one of which was involved in the development of the custom gestures used within the software. Whilst this showed that the software had good agreement compared to raters who closely followed the FMS scoring criteria, it did not provide any indication of the level of agreement with raters with varying levels of experience, which would be more representative of the circumstances found within applied settings. Consideration should also be given to testing the accuracy and reliability of the novel software against alternative, objective, automated measurement systems such as inertial measurement units and marker based motion capture systems. This would establish if the software is accurately tracking key joints and providing precise data related to the movement of these joints during subtest performance compared to these systems. This programme of research was only able to compare the software's performance to subjective manual scoring. As manual scoring is based on each testers interpretation of the FMS scoring criteria, using such a method as the reference against which to compare automated technology, may have led to inaccuracies in the results recorded.

Finally, further validation work in this area should involve a larger sample size, to include participants that have a wide range of body shapes, anthropometric measures and from a variety of sports. The validations to date have not included a large enough sample to be able to establish if large differences between participants have a negative impact on the Kinect sensor's ability to track joint centres. The programme of research was not able to specifically analyse if differences in body shape had any impact on the accuracy or reliability of the software, because of the random nature and sizes of the samples used. For automated software to be confidently used within applied settings, it is important to understand if differences in body shape and size has any impact on the software's performance.

CHAPTER SEVEN

CONCLUSION

7.0 INTRODUCTION

This programme of research has discussed the FMS in great depth and focussed on its ability to predict injury and performance, as well as attempting to improve the reliability of the screening protocol. The FMS is a relatively new tool and is a growing area of interest, as medical and sports science staff attempt to better understand the role fundamental movement can play to improve an individual’s well-being. Quantifying human movement from a multi-joint perspective has become increasingly common (Wainner, Whitman, Cleland, & Flynn, 2007), as the acceptance that compensation in one area of the body could be caused by, or lead to, a weakness in another part of the body, has increased in recent years (Gulgin & Hoogenboom, 2014). The FMS offers a simple and cost-effective diagnostic approach to assess multi-joint movement, which can help guide a practitioner in the best course of action for each individual under their care. However, despite this increased level of acceptance, it is still unclear how to most effectively use the FMS from an applied perspective, and if indeed it is a suitable tool to use in athletic populations. Whilst certain research has suggested it can effectively predict injury rates in different sporting populations (Chorba et al., 2010; Kiesel et al., 2007), there are also studies that contradict these findings (Appel, 2012; Bardenett et al., 2015), suggesting the link between the FMS and injury is still questionable. From an applied perspective, practitioners regularly conduct the FMS as part of a pre-participation screening protocol but are often wary of using it as a standalone measure to predict injury risk. Much of the literature looking at FMS and injury include several limitations regarding study design which restrict the significance of the link between FMS score and injury suggested in these findings. There is also currently a lack of research that has analysed the effectiveness of corrective exercise interventions based on FMS results, that are designed to increase FMS performance, and whether these lead to reduced injury risk.

Chapter three discussed the first study in this programme of research which assessed the effectiveness of the FMS in predicting injury risk and performance in a sample of soccer players. Players undertook the seven FMS subtests and five performance tests during the pre-season period, before their injury occurrence rates were tracked throughout one season. This research used a standardised time loss definition of injury as per the UEFA Injury model (Hagglund et al., 2005), considered a range of injury types and tracked match exposure levels for each player. In addition to using the standard FMS scoring system, the researcher developed a revised FMS scoring system, to attempt to improve mid-range performance assessment for five of the FMS subtests. The study analysed the relationship between total FMS composite scores and injury risk, FMS scores and performance test results and performance test results and injury risk. Consideration was given to the relationship between an FMS cut-off score of 14 or less and different injury types, including non-contact injuries and moderate to severe injuries. In addition, the link between subtest scores and injury risk, along with asymmetries and injury risk, was also analysed to determine if the number of asymmetries impacted a player's injury occurrence.

There has been a range of research studies that have discussed the inter-rater and intra-rater reliability of the FMS. The literature in this area has suggested that the FMS is a reliable screen when considering total FMS composite score, with high levels of inter-rater reliability (Teyhen et al., 2012a) and intra-rater reliability reported (Shultz et al., 2013; Parenteau et al., 2014). However, this level of reliability decreases significantly for subtest scoring (Gulgin & Hoogenboom, 2014; Minick et al., 2010) and this could have serious implications for corrective exercise prescription and how effective these are. The "Corrective Exercise Strategy" recommended in the FMS Manual, strongly indicates that the subtest scores should be used to guide future interventions, rather than the total composite score. Hence, improving the reliability of subtest scoring would improve corrective exercise recommendations and should lead to improvements in an individual's fundamental movement.

Chapters four to six, discussed the development and validation of novel depth camera technology to assess the seven FMS subtests, in an attempt to improve the validity and reliability of subtest scoring compared to manual scoring. Throughout the development process, the researcher and development team attempted to incorporate the scoring criteria outlined in the FMS Manual, to ensure the software replicated manual scoring of the FMS as closely as possible. It was felt that replicating the manual scoring system was important if the system was to be adopted in applied settings. All research relating to the FMS to date, which informs how it is used in practice, was based on the FMS scoring criteria, as were the recommended corrective exercise strategies. Adopting a different scoring system for the automated software would have meant not only validating its accuracy and reliability, but also the new scoring system and corrective exercise prescription process, which was beyond the scope of the programme of research. Each subtest provided its own challenge, as they each had their own scoring parameters that included alternative set up positions to achieve different scores and involved measuring different body parts to assess performance. Following completion of each part of the development process, the software was validated to help inform future development by adjusting the custom gestures set for each subtest, the set-up position and height of the Kinect sensor and the quality and efficiency of the user interface. The final phase of development focussed on all seven subtests as well as allowing the software to automatically prescribe corrective exercises based on the FMS results. The final validation included a comparison of the software's results against both live and video based manual scoring and provided some encouraging results.

7.1 IMPLICATIONS

The findings from chapter three further support the view that the total FMS composite score cannot accurately be used to assess an individual's risk of injury. The study highlighted that a cut-off score of 14 or less did not significantly increase a soccer players risk of injury, compared to those who scored above 14. The same relationship was found for non-contact injuries and for all injury severities, which suggest that the total FMS composite score is not a reliable method to assess injury risk in this sample. Previous research that has discussed FMS and injury, has tended to ignore the impact that exposure levels may have on injury risk (Kiesel et al., 2007), included no definition of injury (Peate et al., 2007) and focussed only on total FMS composite score and injury risk (Garrison et al., 2015; Letafatkar et al., 2014). This has reduced the significance of the findings reported in these studies and meant that the results cannot be used to support the use of the FMS in injury prevention. The current study did not identify any significant increase in injury risk based on FMS score and increased match exposure and could not substantiate the results from previous research that suggest a cut of score of 14 should be used to predict an individual's risk of receiving any type of injury.

Future studies in this area may benefit from analysing the relationship between subtest scores and injury. The current study reported a significant odds ratio of receiving a non-contact injury for players scoring a two or less on the rotary stability subtest, compared to those scoring a three. Although the number of participants scoring a three (n=5) was low in this study, the results suggest further work studying the effect of subtest scores may be a more worthwhile approach than simply considering the total FMS composite score. Linked to this, the effect of corrective exercise interventions, prescribed based on individual subtest scores, would be worthy of future analysis, as to date there is no research that has assessed the effectiveness of such interventions on reducing injury risk.

The study outlined in chapter three found a significant difference in vertical jump height between those scoring above 14 on the FMS and those scoring 14 or below, with the higher FMS group reporting greater jump heights. It also found a significant correlation between vertical jump and Yo-Yo test performance and total FMS score. Although none of the other performance tests reported a significant relationship between performance and FMS scores using a cut-off score of 14, the relationship between the vertical jump and total FMS composite score may be worth further exploration as previous research has also suggested a link between jump height and FMS score (Conlon, 2013). Introducing an intervention group, in a randomised controlled study, should be considered, whereby FMS scores are increased in the intervention group but not the control group, to see if vertical jump height improves in line with increases in FMS score. Increasing sample sizes and including a range of different sports in such a study would also help to understand if there is a link between vertical jump height and total FMS score. Results from the other performance tests support previous research that suggested there is no link between total FMS score and performance (Parchmann & McBride, 2011; Waldron et al., 2014). Given the FMS is designed to assess an individual’s fundamental movement patterns and should be used to detect any pain or weaknesses to direct future treatment or interventions, it is perhaps not surprising that the relationship between total FMS score and performance appears to be weak. With a large amount of variation in the body mechanics involved in different sports, it would be very difficult for a set of seven subtests to reliably predict performance in a multitude of sports.

The results from chapters four to six showed that the software was able to accurately score the hurdle step, shoulder mobility, active straight leg raise and rotary stability subtests compared to live and video based manual scoring, achieving greater agreement levels than that recorded in previous research (Minick et al., 2010; Teyhen et al., 2012). In addition, the test-retest reliability of the software was very high, and once again higher than that recorded by previous research (Shultz et al., 2013; Smith et al., 2013), and the software showed good accuracy compared to manual scoring when prescribing corrective exercises. The software was however, less accurate when scoring the deep squat, inline lunge and trunk stability subtests. This resulted in lower agreement between the software and manual scoring when scoring these three subtests and total FMS composite scores.

The results from the novel software are promising, as they suggest the software has the capability to reduce the subjectivity of FMS scoring and improve the process of accurately prescribing an intervention programme. Implementing such a tool could increase the use of the FMS in applied settings and ensure that practitioners and researchers focus their future work on subtest scores and the interventions prescribed based on these, rather than using total FMS score. The FMS Manual clearly outlines that the subtest results should be carefully analysed to direct future testing or exercise prescription or, in the case of pain being felt, treatment. It suggests that practitioners should primarily focus on mobility weaknesses, followed by stability issues, and finally total body patterning defects, which means the subtest scores are of greater importance than the total FMS composite score. Therefore, a tool that allows the subtests to be scored in an accurate and consistent manner, will provide an opportunity for the FMS to be more effectively utilised in both research and applied settings, and help to identify its most useful application moving forward.

7.2 RECOMMENDATIONS

The findings from this programme of research suggest that the FMS should not be used as a standalone tool to assess injury risk or performance, and that it may be possible to improve the accuracy and reliability of the screening protocol using novel depth camera technology. Based on these findings, the following recommendations can be made regarding the use of the FMS in applied settings and future research in this area.

7.2.1 FMS AND INJURY

As the research relating to the FMS and injury is inconclusive, it would appear that the FMS cannot reliably be used as a standalone test to predict injury risk. However, it may be a more effective tool to use as part of a wider screening protocol, whereby individuals are assessed in a variety of different measures, to provide a holistic overview of an individual’s current level of movement and sport specific competency. This would provide practitioners with benchmark data that can be continuously reviewed and updated as individuals are screened on a regular basis. Therefore, in order to understand the most effective role for the FMS in a sporting environment, additional research is required. This research should follow standardised protocols regarding definition of injury, injury severities and types, inclusion of exposure levels and assessment of subtest scores as well as the total FMS composite score. Until the research in this area can follow such guidelines it will not be possible to understand if the FMS has any link with injury or can be used to accurately predict injury risk.

Previous chapters in this thesis have highlighted the importance of focussing on FMS subtest results rather than the total FMS composite score. These subtest results should guide practitioners as they review results and decide on the appropriate course of action with athletes or individuals under their care following completion of the screening process. Much of the previous research related to the FMS has considered whether a cut-off score of 14 or less can predict an individual's risk of injury or predict levels of performance. However, this may not the most effective use of the FMS in applied and research settings. By analysing subtest scores in more detail, a better understanding of the link between each subtest and injury risk could be established. This would provide a clearer overview of the most effective use of the FMS in applied settings and allow research to be undertaken that can offer critical insight from an injury prevention perspective.

Another area relating to the FMS that has not been well researched to date, is the effectiveness of corrective exercise interventions. The objective of the FMS is to assess an individual's fundamental movement patterns to establish if they have any weaknesses or asymmetries, or indeed, feel pain, when performing each subtest. The results help to guide a practitioner when deciding on the best course of action from an intervention point of view, to ensure an individual's weakest area is the focus of any programme moving forward. Understanding whether these interventions are effective in improving FMS scores and can reduce an individual's risk of injury, would be very worthwhile from both an applied and research perspective. This could be completed using prospective research where FMS scores are increased in an intervention group, compared to a control group, and their injury occurrence rates monitored over a period of time.

Finally, when considering the FMS and injury, future research should understand the physical competencies of their sample and how this may affect the cut-off score used to determine injury risk. As previously discussed in this programme of research, many of the studies to date that have reported a link between FMS score and injury risk, have been conducted on highly physical sports or occupations, which require a high level of physical competence in order to perform at the required level. These studies have reported a mean FMS score that was above the cut-off score of 14 used to identify injury risk. The reverse relationship is apparent for previous studies reporting no link between FMS score and injury, where the mean FMS score was below the cut-off score of 14, suggesting these participants are less physically competent. This suggests that future studies that use a cut-off FMS score to identify injury risk, should consider the sample they are using and set an appropriate cut-off score for that group which reflects the level of physical and movement competency shown by the participants. This cut-off score is likely to be different for each research study but may provide a greater insight into the role FMS can play in identifying injury risk for various sports or occupations.

7.2.2 FMS AND PERFORMANCE

The link between FMS and performance is still inconclusive but is worthy of further consideration. To date the research in this area has used relatively low sample sizes. As this programme of research has identified, the relatively small effect sizes recorded mean large sample sizes are needed to assess if actual significant differences exist between groups with high and low FMS scores. Hence, a sample of approximately 300 participants would be required to ensure such a study is sufficiently powered. Studies that have found a correlation between FMS score and performance have not managed to reach such numbers and as such it is difficult to draw conclusions from the work that has been completed to date.

As discussed earlier in this thesis, to fully understand the relationship between FMS scores and performance, an intervention based study is required. To date, research that has considered whether FMS score impacts performance has used a cross sectional study design and simply screened participants, prior to observing performance over a certain period of time. No attempts have been made to change FMS score and assess the impact on performance using a randomised controlled protocol. FMS score could be improved in an intervention group over a number of weeks and then a reassessment conducted, to determine if their improved movement competency has led to performance enhancements compared to a control group. With so many variables affecting performance, the research to date does not provide sufficient evidence that FMS score does effect performance due to the study designs employed.

7.2.3 AUTOMATED APPROACH

This programme of research has outlined the development and validation process adopted in order to automate the scoring of the FMS. Whilst the automated software was able to reach good levels of reliability and validity compared to manual scoring, there are areas that could be improved if the researcher was able to complete the process again. One of the main changes that should be considered is trying to replicate the existing FMS manual scoring system when designing the automated software. This route was chosen to make the software as relevant as possible from an applied perspective but resulted in the software not being as effective or as automated as had been expected. A more prudent approach may have been to develop a new scoring system, specifically designed to suit the needs of the automated software, which could provide the user with raw data relating to joint positioning, angles and alignment when performing the FMS. This would have removed the need for the custom gestures developed in the software to be matched with the four different scoring options outlined in the FMS scoring criteria. Whilst this would make comparison with manual scoring difficult, it would provide the opportunity to compare against gold standard methods such as 3D motion capture systems, thus providing a clear indication of the level of error within the software when calculating and tracking joint centres. Future research in this area should consider developing an automated system that can be used to support manual scoring, rather than as a direct replacement for it, through the identification and tracking of key joints to support the traditional FMS score provided manually.

In regard to the validation of the software, future research in this area should consider the ability of the individual participants, to ensure a wider spread of scores on each of the FMS subtests compared to this programme of research. A major challenge of the current research was to ensure that each subtest had a similar number of performances that could be scored as a one, two or a three across the validation studies. In most cases the number of performances scored as a one, significantly outweighed those scored as a three. Whilst this is representative of the wider population as a whole in terms of FMS performance, it may have led to higher levels of agreement between manual and automated scoring methods. Using the FMS scoring criteria, a one is typically easier to identify on each subtest due to the relatively obvious compensations being made. For example, on the trunk stability subtest a one is scored if the participant is unable to perform one push up from either of the two starting positions adopted. Therefore, higher levels of agreement between the two scoring methods could have been reported because of the number of low scores given for performances on many of the subtests. The focus of future validation studies in this area should be on achieving an even spread of scores for participants across all seven subtests, even if this requires manipulation of performances to ensure this broad range is reached. This will provide a clearer reflection of the automated systems validity and reliability as it will be required to accurately and consistently score every type of performance across all seven subtests.

APPENDIX A

INFORMED CONSENT FORMS

APPENDIX A.1 – INFORMED CONSENT FORM STUDY ONE

WATERFORD INSTITUTE OF TECHNOLOGY

RESEARCH - INFORMED CONSENT FORM

I. Project Title:

Effectiveness of the Functional Movement Screen in predicting injury rates amongst soccer players

II. Introduction to this study:

The Functional Movement Screen (FMS) is a series of seven subtests that analyse a player's movement patterns and in particular their flexibility, mobility, and stability across several parts of the body. These tests are increasingly being used by sports teams to help coaches develop corrective exercise programmes for players to reduce their injury risk. Several studies have shown the effectiveness of the FMS in predicting injury risk amongst different populations. With soccer being a multi directional field based sport, there are alternative soccer specific tests that may provide a more accurate prediction of injury amongst soccer players than FMS subtests in isolation.

III. I am being asked to participate in this research study. The study has the following purposes:

1. To determine how effective Functional Movement Screening (FMS) is in predicting injury amongst soccer players compared to alternative screening methods
2. To analyse each of the seven FMS tests accuracy in predicting injury in soccer players in an attempt to determine if certain tests are more relevant than others
3. To determine the correlation between FMS tests and soccer specific tests to ascertain if FMS scores provide an indication of soccer performance
4. To analysis if correlations exist in regard to FMS scores, player ages, player position and injury prediction accuracy

IV. This research study will take place at your club's training ground

V. This is what will happen during the research study:

1. You will be asked to complete a Pre-Test Questionnaire detailing your name, age and previous medical history. You will also be weighed and have your height measured for same

2. You will then participate in the seven FMS Tests, where the researcher will provide full instruction. You will be given a score for your performance on each test

3. You will then participate in the seven-soccer specific/functional tests, where the researcher will provide full instruction. You will be given a score for your performance on each test

4. During the 2014 and 2015 season your Club Physiotherapist will track your injury occurrence rates in both training and match situations, using our Injury Data Collection Sheets

VI. There are minimal risks or side effects associated with participation in this study. These are:

Participation in strenuous exercise could trigger a heart attack in individuals with an underlying heart condition

VII. There may be benefits from my participation in this study. These are:

The research will help with the development of more appropriate screening protocols which may assist in injury reduction for soccer players in the future

VI. My confidentiality will be guarded:

Waterford Institute of Technology will protect all the information about me and my part in this study. My identity or personal information, will not be revealed, published or used in future studies. The study findings will form the basis for preparation of a postgraduate thesis, academic publications, conference papers and other scientific publications.

VII. If I have questions about the research project, I am free to call Dr. Michael Hanlon at telephone no. 051 302166

VIII. Taking part in this study is my decision.

If I do agree to take part in the study, I may withdraw at any point. There will be no penalty if I withdraw before I have completed all stages of the study.

IX. Signature:

I have read and understood the information in this form. My questions and concerns have been answered by the researchers, and I have a copy of this consent form. Therefore, I consent to take part in this research project entitled: Effectiveness of the Functional Movement Screen in predicting injury rates amongst soccer players.

Signed: ______________________________

Print Name: ……………………………………………………………………

Date: ______________________________

Witness: ______________________________

Signature: ______________________________

APPENDIX A.2 – INFORMED CONSENT FORM STUDY TWO

WATERFORD INSTITUTE OF TECHNOLOGY

RESEARCH - INFORMED CONSENT FORM

I. <u>Project Title</u>: Development of motion tracking software to assess the deep squat FMS subtest

II. <u>Introduction to this study:</u>

The Functional Movement Screen (FMS) is a series of seven subtests that analyse an individual's movement patterns and in particular their flexibility, mobility, and stability across several parts of the body. These subtests are increasingly being used by sports teams to help coaches develop corrective exercise programmes for players to reduce their injury risk. Several studies have shown the effectiveness of FMS in predicting injury risk amongst different populations. The researcher has developed innovative software that allows the deep squat FMS subtest to be analysed automatically without the need for manual assessment.

III. <u>Participants are being asked to participate in this research study. The study has the following purposes:</u>

1. Can software, integrated with the Kinect sensor v2, be developed that can validly assess performance on the deep squat FMS subtest compared to blinded video scoring by a certified FMS tester?
2. Can software, integrated with the Kinect sensor v2, be developed that increases the test retest reliability of the deep squat FMS subtest compared to manual screening by a certified FMS tester?

IV. <u>This research study will take place at the GSK Human Performance Lab, Great West Road, London</u>

V. <u>This is what will happen during the research study:</u>

1. Participants will be asked to complete a Pre-Test Questionnaire detailing their name, age and previous medical history.

2. Participants will then participate in the deep squat FMS subtest, where the researcher will provide full instruction. Participants will be given a score for their performance on the subtest by the researcher as well as from the software that will automatically assess their performance.

3. Participants will be requested to return to the Lab seven days after the first testing session to repeat the same deep squat protocol.

VI. There are minimal risks or side effects associated with participation in this study. These are:

Participation in the deep squat FMS subtest could cause acute pain which could make performance difficult

VII. There may be benefits from my participation in this study. These are:

The research will help with the development of the software to ensure it is as accurate and effective as possible, and be able to assess all seven FMS tests.

VI. Participant's confidentiality will be guarded:

Waterford Institute of Technology will protect all the information about participant's and their part in this study. Their identity or personal information will not be revealed, published or used in future studies. The study findings will form the basis for preparation of a postgraduate thesis, academic publications, conference papers and other scientific publications.

VII. If participant's have questions about the research project, they are free to call Dr. Michael Hanlon at telephone no. 051 302166

VIII. Taking part in this study is the participant's decision.

If participants do agree to take part in the study, they may withdraw at any point. There will be no penalty if they withdraw before they have completed all stages of the study.

IX. Signature:

I have read and understood the information in this form. My questions and concerns have been answered by the researchers, and I have a copy of this consent form. Therefore, I consent to taking part in this research project entitled: Development of motion tracking software to assess the deep squat FMS subtest.

Signed: ____

Print Name: ……………………………………………………………………………

Date: ____

Witness: ____

APPENDIX A.3 – INFORMED CONSENT FORM STUDY THREE

WATERFORD INSTITUTE OF TECHNOLOGY

RESEARCH - INFORMED CONSENT FORM

I. Project Title: Development of motion tracking software to assess the three FMS subtests

II. Introduction to this study:

The Functional Movement Screen (FMS) is a series of seven subtests that analyse an individual's movement patterns and in particular their flexibility, mobility, and stability across several parts of the body. These subtests are increasingly being used by sports teams to help coaches develop corrective exercise programmes for players to reduce their injury risk. Several studies have shown the effectiveness of FMS in predicting injury risk amongst different populations. The researcher has developed innovative software that allows the deep squat, hurdle step and inline lunge FMS subtests to be analysed automatically without the need for manual assessment.

III. Participants are being asked to participate in this research study. The study has the following purpose:

Can software, integrated with the Kinect sensor v2, be developed that can validly and reliably assess performance on the deep squat, hurdle step and inline lunge FMS subtests compared to manual scoring by a certified FMS tester?

IV. This research study will take place at your clubs training ground

V. This is what will happen during the research study:

1. Participants will be asked to complete a Pre-Test Questionnaire detailing their name, age and previous medical history.

2. Participants will then participate in the deep squat, hurdle step and inline lunge FMS subtests, where the researcher will provide full instruction. Participants will be given a score for their performance on each subtest by the researcher as well as from the software that will automatically assess their performance.

VI. There are minimal risks or side effects associated with participation in this study. These are:

Participation in the FMS subtests could cause acute pain which could make performance difficult

APPENDIX A.3 (CONT) – INFORMED CONSENT FORM STUDY THREE

VII. There may be benefits from my participation in this study. These are:

The research will help with the development of the software to ensure it is as accurate and effective as possible, and be able to assess all seven FMS tests.

VI. Participant's confidentiality will be guarded:

Waterford Institute of Technology will protect all the information about participant's and their part in this study. Their identity or personal information will not be revealed, published or used in future studies. The study findings will form the basis for preparation of a postgraduate thesis, academic publications, conference papers and other scientific publications.

VII. If participant's have questions about the research project, they are free to call Dr. Michael Hanlon at telephone no. 051 302166

VIII. Taking part in this study is the participant's decision.

If participants do agree to take part in the study, they may withdraw at any point. There will be no penalty if they withdraw before they have completed all stages of the study.

IX. Signature:

I have read and understood the information in this form. My questions and concerns have been answered by the researchers, and I have a copy of this consent form. Therefore, I consent to taking part in this research project entitled: Development of motion tracking software to assess the three FMS subtests.

Signed: ____

Print Name: ………………………………………………………………………………

Date: ____

Witness: ____

APPENDIX A.4 – INFORMED CONSENT FORM STUDY FOUR

WATERFORD INSTITUTE OF TECHNOLOGY

RESEARCH - INFORMED CONSENT FORM

I. Project Title: Development of motion tracking software to assess the seven FMS subtests

II. Introduction to this study:

The Functional Movement Screen (FMS) is a series of seven subtests that analyse an individual's movement patterns and in particular their flexibility, mobility, and stability across several parts of the body. These tests are increasingly being used by sports teams to help coaches develop corrective exercise programmes for players to reduce their injury risk. Several studies have shown the effectiveness of FMS in predicting injury risk amongst different populations. The researcher has developed innovative software that allows the FMS subtests to be analysed automatically without the need for manual assessment.

III. Participants are being asked to participate in this research study. The study has the following purposes:

1. To assess if the technology developed provides accurate assessment of performance on the FMS and increases the inter-rater and test-retest reliability of the FMS compared to manual testing by an expert tester
2. To assess if the technology developed can automatically assess the FMS, and prescribe a suitable intervention program compared to manual testing?

IV. This research study will take place at Drogheda Institute of Further Education, The Twenties, Drogheda, Louth

V. This is what will happen during the research study:

1. Participants will be asked to complete a Pre-Test Questionnaire detailing their name, age and previous medical history.

2. Participants will have their height and weight measurements taken

3. Participants will then participate in seven FMS subtests (deep squat, hurdle step, inline lunge, shoulder mobility, active straight leg raise, trunk stability push up and rotary stability), where the researcher will provide full instruction. Participants will be given a score for their performance on each subtest by the researcher as well as from the software that will automatically assess their performance.

4. Exactly one-week later participants will perform each FMS subtest again in the same order as outlined above. The researcher will provide full instruction. Participants will be given a score for their performance on each subtest by the researcher as well as from the software that will automatically assess their performance

VI. There are minimal risks or side effects associated with participation in this study. These are:

Participation in the FMS could cause acute pain which could make performance difficult

VII. There may be benefits from my participation in this study. These are:

The research will help with the development of the software to ensure it is as accurate and effective as possible, and be able to assess all seven FMS subtests.

VI. Participant's confidentiality will be guarded:

Waterford Institute of Technology will protect all the information about participant's and their part in this study. Their identity or personal information will not be revealed, published or used in future studies. The study findings will form the basis for preparation of a postgraduate thesis, academic publications, conference papers and other scientific publications.

VII. If participant's have questions about the research project, they are free to call Dr. Michael Hanlon at telephone no. 051 302166

VIII. Taking part in this study is the participant's decision.

If participants do agree to take part in the study, they may withdraw at any point. There will be no penalty if they withdraw before they have completed all stages of the study.

IX. Signature:

I have read and understood the information in this form. My questions and concerns have been answered by the researchers, and I have a copy of this consent form. Therefore, I consent to taking part in this research project entitled: Improving the reliability of the Functional Movement Screen through the development of motion tracking technology

Signed: ____

Print Name: ………………………………………………………………………………

Date: ____

APPENDIX B

PRE-TEST QUESTIONNAIRES

APPENDIX B.1 – PRE-TEST QUESTIONNAIRE STUDY ONE

PRE-TEST QUESTIONNAIRE	
NAME (PLEASE PRINT)	
DATE OF BIRTH	
PLAYING POSITION	
DO YOU CURRENTLY HAVE AN INJURY THAT PREVENTS YOU FROM PARTICIPATING IN SOCCER ACTIVITIES?	
IF YES, PLEASE STATE TYPE OF INJURY AND LENGTH OF ABSENCE FROM SOCCER ACTIVITIES	
HAVE YOU HAD AN INJURY IN THE PAST 5 YEARS THAT HAS RESULTED IN YOU BEING UNABLE TO PARTICIPATE IN SOCCER ACTIVITIES FOR 3 WEEKS OR MORE?	
IF YES PLEASE STATE TYPE OF INJURY AND LENGTH OF ABSENCE FROM SOCCER ACTIVITIES (INCLUDE MORE THAN ONE INJURY IF RELEVANT)	1 2 3
HAVE YOU HAD AN INJURY THAT HAS REQUIRED SURGERY AT ANY POINT DURING YOUR CAREER?	
IF YES PLEASE STATE INJURY TYPE	1 2
TO BE COMPLETED BY RESEARCHER	
HEIGHT AND WEIGHT	
DOMINANT FOOT	

APPENDIX B.2 – PRE-TEST QUESTIONNAIRE STUDY TWO

PRE-TEST QUESTIONNAIRE	
NAME (PLEASE PRINT)	
DATE OF BIRTH	
PLEASE LIST ANY FORM OF EXERCISE THAT YOU PARTICIPATE IN THREE OR MORE TIMES PER WEEK	
DO YOU CURRENTLY HAVE AN INJURY THAT PREVENTS YOU FROM PARTICIPATING IN ANY FORM OF EXERCISE?	
IF YES, PLEASE STATE TYPE OF INJURY AND LENGTH OF ABSENCE FROM ACTIVITIES	
HAVE YOU HAD AN INJURY IN THE PAST 5 YEARS THAT HAS RESULTED IN YOU BEING UNABLE TO PARTICIPATE IN EXERCISE FOR 3 WEEKS OR MORE?	
IF YES PLEASE STATE TYPE OF INJURY AND LENGTH OF ABSENCE FROM ACTIVITIES **(INCLUDE MORE THAN ONE INJURY IF RELEVANT)**	1 2 3
HAVE YOU HAD AN INJURY THAT HAS REQUIRED SURGERY **AT ANY POINT IN YOUR LIFE?**	
IF YES PLEASE STATE INJURY TYPE	1 2
TO BE COMPLETED BY RESEARCHER	
HEIGHT AND WEIGHT	
DOMINANT HAND/FOOT	

APPENDIX B.3 – PRE-TEST QUESTIONNAIRE STUDY THREE & FOUR

PRE-TEST QUESTIONNAIRE	
NAME (PLEASE PRINT)	
DATE OF BIRTH	
CHOSEN SPORT AND PLAYING POSITION	
DO YOU CURRENTLY HAVE AN INJURY THAT PREVENTS YOU FROM PARTICIPATING IN YOUR CHOSEN SPORT?	
IF YES, PLEASE STATE TYPE OF INJURY AND LENGTH OF ABSENCE FROM YOUR CHOSEN SPORT	
HAVE YOU HAD AN INJURY IN THE PAST 5 YEARS THAT HAS RESULTED IN YOU BEING UNABLE TO PARTICIPATE IN YOUR CHOSEN SPORT FOR 3 WEEKS OR MORE?	
IF YES PLEASE STATE TYPE OF INJURY AND LENGTH OF ABSENCE FROM YOUR CHOSEN SPORT **(INCLUDE MORE THAN ONE INJURY IF RELEVANT)**	1 2 3
HAVE YOU HAD AN INJURY THAT HAS REQUIRED SURGERY AT ANY POINT IN YOUR LIFE?	
IF YES PLEASE STATE INJURY TYPE	1 2
TO BE COMPLETED BY RESEARCHER	
HEIGHT AND WEIGHT	
DOMINANT HAND/FOOT	

APPENDIX C

STUDY ONE DATA COLLECTION

APPENDIX C.1 – THE 16 LOWER EXTREMITY FUNCTIONAL TEST SHUTTLES

Shuttle	Direction	Skill Description	Sequence
1	Forward	Forward sprint between targets, turning at C	A, C, A
2	Backward	Backward sprint between targets, turning at C	A, C, A
3	Right	Side steps to the right around entire course facing the centre	A, D, C, B, A
4	Left	Side steps to the left around entire course facing the centre	A, B, C, D, A
5	Right	Carioca to the right around entire course facing the centre	A, D, C, B, A
6	Left	Carioca to the right around entire course facing the centre	A, B, C, D, A
7	Right	Figure of 8 sprint running around D and B and circling C and A inside out	A, D, C, B, A
8	Left	Figure of 8 sprint running around D and B and circling C and A inside out	A, B, C, D, A
9	Right	Forward run to right plant outside foot for 45° cut to next target	A, D, C, B, A
10	Left	Forward run to left plant outside foot for 45° cut to next target	A, B, C, D, A
11	Right	Forward run to right plant outside foot for 90° cut to next target	A, D, B, A
12	Left	Forward run to left plant outside foot for 90° cut to next target	A, B, D, A
13	Right	Forward run to right plant inside foot to cross outside foot for 90° cut to next target	A, D, B, A
14	Left	Forward run to left plant inside foot to cross outside foot for 90° cut to next target	A, B, D, A
15	Forward	Forward sprint between targets, turning at C	A, C, A
16	Backward	Backward sprint between targets, turning at C	A, C, A

APPENDIX C.2 – SCREENING TEST SCORE SHEET STUDY ONE

Name of Participant

Test	Raw Score	Final Score	Revised FMS Score
Deep Squat			
Hurdle Step			
Left Leg			
Right Leg			
In line Linge			
Left Leg			
Right leg			
Shoulder Mobility			
Left Shoulder			
Right Shoulder			
Active Straight Leg Raise			
Left Leg			
Right Leg			
Trunk Stability Push Up			
Rotary Stability			
Left Side			
Right Side			

Total Score

APPENDIX C.2 (CONT) – SCREENING TEST SCORE SHEET STUDY ONE

Name of Participant	

Single Leg Hop Test

Attempts	1	2
Right Leg		
Left Leg		

Drop Jump Test

	Knee-Knee	Ankle-Ankle	Difference
Jump 1			
Jump 2			
Jump 3			

Vertical Jump

	Height (cm)
Jump 1	
Jump 2	
Jump 3	

Lower Extremity Functional Test

Time in Seconds	

Yo-Yo Recovery Level 2 Intermittent Test

Shuttle	
Level	
Metres	

APPENDIX C.3 – INJURY DATA COLLECTION SHEET

Injury No	Player No	When did Injury occur	Date injury occurred	Location of injury	Nature of injury	Mechanism of injury	Period of match/training injury occurred	Estimated severity of injury	Has the player returned to full training	If yes, what date did they return	Total days absence	How many match/training days were missed	Was this a reinjury

APPENDIX C.3 (CONT) – INJURY DATA COLLECTION SHEET

Definitions of terms used in Injury Data Collection sheet

1. Match
Any organised club game lasting ninety minutes (or 120 minutes in the case of extra time)

2. Training
A coach directed physical activity carried out with a team

3. Injury
A recordable injury sustained during a match or training session inside club that results in absence from soccer participation

4. Actual number of days lost due to injury
Count number of days between injury occurring and player returning to full training

5. Full training
A player is only considered full rehabilitated when they can participate in 100% of the training session

6. Reinjury
An injury of the same type and at the same site as a previous injury. Categorised as follows:

Less than 2 months = Early Recurrence
2 to 12 months = Late Recurrence
More than 12 months = Delayed Recurrence

7. injury Severity is defined as follows:

Slight (1 - 3 days absence)
Minor (4 - 7 days absence)
Moderate (8 - 28 days absence)
Major (More than 28 days absence)

APPENDIX C.4 – TRAINING MINUTES FORM

Week No	Week Commencing	Number of training minutes	Number of players who completed all training minutes

APPENDIX C.5 – MATCH MINUTES FORM

Match Minutes

Player Number	Match 1 Mins	Match 2 Mins	Match 3 Mins	Match 4 Mins	Match 5 Mins	Match 6 Mins	Match 7 Mins	Match 8 Mins	Match 9 Mins	Match 10 Mins	Match 11 Mins	Total Season Mins
												0
												0
												0
												0
												0
												0
												0
												0
												0
												0
												0
												0
												0
												0
												0
												0
												0
												0
												0

APPENDIX D

STUDY ONE RESULTS

APPENDIX D.1 - SUMMARY OF FMS CUT-OFF SCORES, SENSITIVITY AND SPECIFICITY, AND THE ODDS RATIO OF RECEIVING ANY TYPE OF INJURY AND A NON-CONTACT INJURY

All injury types

FMS Score	Injured (n)	Not Injured (n)	Odds Ratio	95% CI	*p*	Sensitivity (%)	Specificity (%)
≤13	4	0	N/A	N/A	N/A	7	100
≥14	53	32					
≤14	9	5	1.01	0.31-3.33	1	16	84
≥15	48	27					
≤15	19	9	1.28	0.50-3.29	0.6	33	72
≥16	38	23					
≤16	27	15	1.02	0.43-2.43	1	47	53
≥17	30	17					
≤17	39	27	0.4	0.13-1.21	0.1	68	16
≥18	18	5					
≤18	51	30	0.58	0.11-2.99	0.5	89	6
≥19	6	3					

Non-contact injuries

FMS Score	Injured (n)	Not Injured (n)	Odds Ratio	95% CI	*p*	Sensitivity (%)	Specificity (%)
≤13	1	3	0.39	0.04-3.93	0.41	3	94
≥14	39	46					
≤14	5	9	0.63	0.19-2.07	0.45	13	82
≥15	35	40					
≤15	14	14	1.35	0.55-3.30	0.52	35	71
≥16	26	35					
≤16	21	21	1.47	0.64-3.41	0.37	53	57
≥17	19	28					
≤17	29	37	0.86	0.33-2.21	0.75	73	24
≥18	11	12					
≤18	37	44	1.4	0.31-6.26	0.66	93	10
≥19	3	5					

FMS = Functional Movement Screen; CI = Confidence interval

APPENDIX D.2 - SUMMARY OF FMS CUT-OFF SCORES, SENSITIVITY AND SPECIFICITY, AND THE ODDS RATIO OF BEING ABSENT FOR 8 DAYS OR MORE

			8 Days absence or more				
FMS Score	**8 days or more absence (n)**	**Less than 8 days absence (n)**	**Odds Ratio**	**95% CI**	***p***	**Sensitivity (%)**	**Specificity (%)**
≤13	4	0	N/A	N/A	N/A	10	100
≥14	35	50					
≤14	8	6	1.89	0.60-6.00	0.27	21	88
≥15	31	44					
≤15	14	13	1.44	0.58-3.60	0.43	36	72
≥16	25	36					
≤16	21	21	1.61	0.69-3.75	0.27	54	58
≥17	18	29					
≤17	28	28	1.09	0.41-2.88	0.65	72	30
≥18	11	12					
≤18	35	46	0.76	0.18-3.26	0.71	90	8
≥19	4	4					

FMS = Functional Movement Screen; CI = Confidence interval

APPENDIX D.3 - SUMMARY OF FMS CUT-OFF SCORES USING THE REVISED FMS SCORING SYSTEM, SENSITIVITY AND SPECIFICITY, AND THE ODDS RATIO OF RECEIVING ANY TYPE OF INJURY AND A NON-CONTACT INJURY

All injury types								Non-contact injuries							
FMS Score	**Injured (n)**	**Not Injured (n)**	**Odds Ratio**	**95% CI**	***p***	**Sensitivity (%)**	**Specificity (%)**	**FMS Score**	**Injured (n)**	**Not Injured (n)**	**Odds Ratio**	**95% CI**	***p***	**Sensitivity (%)**	**Specificity (%)**
≤14	6	3	1.14	0.26-4.89	0.86	11	91	≤14	2	7	0.32	0.06-1.61	0.15	27	81
≥15	51	29						≥15	38	42					
≤15	10	5	1.15	0.36-3.71	0.82	18	84	≤15	6	9	0.78	0.25-2.43	0.67	15	82
≥16	47	27						≥16	34	40					
≤16	19	8	1.5	0.57-3.96	0.41	33	75	≤16	15	12	1.85	0.74-4.61	0.18	38	76
≥17	38	24						≥17	25	37					
≤17	25	13	1.14	0.47-2.75	0.77	44	59	≤17	20	18	1.72	0.74-4.03	0.21	50	63
≥18	32	19						≥18	20	31					
≤18	32	21	0.67	0.27-1.65	0.38	56	34	≤18	25	28	1.25	0.53-2.94	0.61	63	43
≥19	25	11						≥19	15	21					
≤19	45	29	0.39	0.10-1.49	0.16	79	9	≤19	33	41	1.03	0.35-3.07	0.88	83	18
≥20	12	3						≥20	7	9					
≤20	48	30	0.41	0.07-1.76	0.19	79	10	≤20	35	43	0.98	0.27-3.47	0.97	88	12
≥21	9	2						≥21	7	9					

FMS = Functional Movement Screen; CI = Confidence interval

APPENDIX D.4 - SUMMARY OF FMS CUT-OFF SCORES USING THE REVISED SCORING SYSTEM, SENSITIVITY AND SPECIFICITY, AND THE ODDS RATIO OF BEING ABSENT FOR 8 DAYS OR MORE

8 Day absence or more using revised FMS scoring system							
FMS Score	**8 days or more absence (n)**	**Less than 8 days absence (n)**	**Odds Ratio**	**95% CI**	***p***	**Sensitivity (%)**	**Specificity (%)**
≤14	5	4	1.69	0.42-6.77	0.45	13	92
≥15	34	46					
≤15	9	6	2.2	0.71-6.83	0.17	23	88
≥16	30	44					
≤16	15	12	1.98	0.79-4.94	0.14	38	76
≥17	24	38					
≤17	18	20	1.29	0.55-3.00	0.56	46	60
≥18	21	30					
≤18	23	30	0.96	0.41-2.25	0.92	59	40
≥19	16	20					
≤19	32	42	0.87	0.29-2.65	0.8	82	16
≥20	7	8					
≤20	32	46	0.4	0.11-1.47	0.16	82	8
≥21	7	4					

FMS = Functional Movement Screen; CI = Confidence interval

APPENDIX D.5 - SUMMARY OF SUBTEST SCORES, SENSITIVITY AND SPECIFICITY, AND THE ODDS RATIO OF RECEIVING ANY TYPE OF INJURY AND A NON-CONTACT INJURY

All Injury Types							Non-Contact Injuries						
Subtest Score	**Injured (n)**	**Not Injured (n)**	**Odds Ratio**	***p***	**Sensitivity (%)**	**Specificity (%)**	**Subtest Score**	**Injured (n)**	**Not Injured (n)**	**Odds Ratio**	***p***	**Sensitivity (%)**	**Specificity (%)**
DS ≤ 2	48	27	0.99	0.97	64	36	DS ≤ 2	33	42	0.79	0.89	44	50
DS 3	9	5					DS 3	7	7				
HS ≤ 2	52	29	1.08	0.7	64	38	HS ≤ 2	36	45	0.8	0.88	44	50
HS 3	5	3					HS 3	4	4				
IL ≤ 2	29	21	0.54	0.38	58	28	IL ≤ 2	20	30	0.63	0.55	40	49
IL 3	28	11					IL 3	20	19				
SM ≤ 2	24	13	1.06	0.46	65	37	SM ≤ 2	19	18	1.56	0.19	51	60
SM 3	33	19					SM 3	21	31				
ASLR ≤ 2	29	14	1.33	0.81	67	39	ASLR ≤ 2	22	21	1.63	0.42	51	61
ASLR 3	28	18					ASLR 3	18	28				
TS ≤ 2	4	3	0.73	0.69	57	35	TS ≤ 2	1	6	0.18	0.09	14	52
TS 3	53	29					TS 3	39	43				
RS ≤ 2	55	29	2.84	0.3	65	60	RS ≤ 2	40	44	3.64	0.05	47	83
RS 3	2	3					RS 3	1	4				

APPENDIX D.6 - SUMMARY OF MEAN PERFORMANCE TEST SCORES, SENSITIVITY AND SPECIFICITY, AND THE ODDS RATIO OF RECEIVING ANY TYPE OF INJURY AND A NON-CONTACT INJURY

All injury types								
Tests	**Cut-off**	**Injured (n)**	**Not Injured (n)**	**Odds Ratio**	**95% CI**	***p***	**Sensitivity (%)**	**Specificity (%)**
VJ	≤51.7 cm	25	15	0.87	0.33-2.30	0.78	54	42
	≥51.8cm	21	11					
LEFT	≤110.54 secs	9	5	1.54	0.33-7.23	0.58	56	55
	≥110.55 secs	7	6					
Yo-Yo	≥1377 m	19	6	2.74	0.84-8.94	0.09	56	68
	≤1376 m	15	13					
SLH Lt Leg	≤75% of height	11	8	0.76	0.23-2.44	0.64	35	58
	≥76% of height	20	11					
SLH Rt Leg	≤75% of height	15	5	2.63	0.76-9.08	0.12	48	74
	≥76% of height	16	14					
DJ	≥2.7cm	26	16	0.91	0.39-3.06	0.86	38	64
	≤2.6cm	16	9					

Non-contact injuries								
Tests	**Cut-off**	**Injured (n)**	**Not Injured (n)**	**Odds Ratio**	**95% CI**	***p***	**Sensitivity (%)**	**Specificity (%)**
VJ	≤51.7 cm	16	24	0.52	0.20-1.33	0.17	47	37
	≥51.8cm	18	14					
LEFT	≤110.54 secs	6	8	1.2	0.26-5.59	0.81	55	50
	≥110.55 secs	5	8					
Yo-Yo	≥1377 m	14	11	1.97	0.66-5.88	0.22	56	61
	≤1376 m	11	17					
SLH Lt Leg	≤75% of height	8	11	0.67	0.25-2.45	0.67	35	59
	≥76% of height	15	16					
SLH Rt Leg	≤75% of height	12	8	2.59	0.81-8.29	0.11	52	70
	≥76% of height	11	19					
DJ	≥2.7cm	19	23	1.05	0.25-2.58	0.92	37	62
	≤2.6cm	11	14					

VJ = Vertical Jump Test; LEFT = Lower Extremity Functional Test; Yo-Yo = Yo-Yo intermittent Recovery Test level 2; SLH = Single Leg Hop Test; DJ = Drop Jump Test

APPENDIX D.7 – EFFECT SIZE AND POWER CALCULATIONS FOR STUDY ONE TESTS

FMS and absences of 8 days or more

	Mean	SD	N
Injured	16.1	2.1	39
Not injured	16.5	1.5	50

SD pooled	1.787022
Effect Size	-0.22
Group size (n)	326

SD Pooled Calculation = SQRT(((2.1^2*(39-1)+1.5^2*(50-1)))/(39+50-2))

Effect Size Calculation = (16.1-16.5)/1.787022

Inline lunge and any type of injury

	Mean	SD	N
Injured	2.45	0.57	57
Not injured	2.31	0.54	32

SD pooled	0.559495
Effect Size	0.25
Group size (n)	253

SD Pooled Calculation = SQRT(((0.57^2*(57-1)+0.54^2*(32-1)))/(57+32-2))

Effect Size Calculation = (2.45-2.31)/0.559495

Rotary stability and any type of injury

	Mean	SD	N
Injured	2	0.27	57
Not injured	2.09	0.3	32

SD pooled	0.281057
Effect Size	-0.32
Group size (n)	155

SD Pooled Calculation = SQRT(((0.27^2*(57-1)+0.3^2*(32-1)))/(57+32-2))

Effect Size Calculation = (2-2.09)/0.281057

APPENDIX D.7 (CONT) – EFFECT SIZE AND POWER CALCULATIONS FOR STUDY ONE TESTS

LEFT and any type of injury

	Mean	SD	N
Injured	111.03	4.79	16
Not injured	109.83	5.85	11

SD pooled	5.239796
Effect Size	0.23
Group size (n)	298

SD Pooled Calculation = SQRT(((4.79^2*(16-1)+5.85^2*(11-1)))/(16+11-2))

Effect Size Calculation = (111.03-109.83)/5.239796

APPENDIX E

PROTOTYPE V1 SCORESHEET

APPENDIX E.1 – PROTOTYPE V1 TESTING SCORE SHEET

Participant No	DS Score by Prototype	DS Score Manual	Left Shoulder°	Right Shoulder°	Left Hip°	Right Hip°	Left Knee°	Left Ankle°	Comments

APPENDIX F

FMS SUBTEST DESCRIPTIONS

APPENDIX F.1 – DESCRIPTION OF DEEP SQUAT & HURDLE STEP

Deep Squat

This subtest challenges total body movement. It is used to assess bilateral and functional mobility of the hips, knees, and ankles and also assesses bilateral and symmetrical mobility of the shoulders, as well as the thoracic spine, with the dowel being placed overhead. The participant places their feet shoulder width apart with the feet aligned in the sagittal plane. The tester should check for correct placement of the feet by placing the dowel horizontally on the ground next to the inside of one foot, to check it resides inline with the crease of the shoulder joint. This should be repeated for the other foot. The individual then adjusts their hands on the dowel to assume a 90° angle of the elbows with the dowel overhead. The tester should guide the participant to ensure the hands are positioned correctly prior to the performance starting. The dowel is then pressed overhead with the shoulders flexed and abducted, and the elbows extended. The individual is then instructed to descend slowly into a squat position. The heels of both feet should remain on the floor during squat performance, with the head and chest facing forward, knees tracking over the feet in the frontal plane and the dowel maximally pressed overhead at all times. If the criteria for a score of three is not achieved, the athlete is then asked to perform the test with the FMS board under their heels. The participant can repeat the test up to three times.

Hurdle Step

The hurdle step subtest is designed to test an individual’s stepping motion. Co-ordination is required between the lower and upper body, whilst single leg stability is also tested. When performing the hurdle step subtest an individual is required to raise one leg over a rubber band, set at the same height as their tibial tuberosity, before stepping forward with the same leg so the heel of that foot touches the ground. The individual is then required to return the foot to the start position by repeating the action in reverse. The entire action is completed whilst the 4ft dowel is held on the individual’s shoulders and with minimal movement of the trunk and standing leg. The individual must keep their head up with eyes looking forward at all times during performance. The test is completed three times on each leg.

APPENDIX F.2 – DESCRIPTION OF INLINE LUNGE & SHOULDER MOBILITY

Inline Lunge

The inline lunge subtest places the feet in a scissor position and tests the body's ability to maintain balance, resist rotation and maintain alignment. It provides an assessment of hip and ankle mobility, and knee stability. To perform the inline lunge a participant places the toes of one foot in line with the red line on the FMS 6ft by 2ft board. The heel of the other foot is then placed in front of that foot the same distance apart as the tibial tuberosity measurement from the hurdle step subtest. The participant then places the 4ft dowel behind their back and holds it vertically with one hand placed behind their neck and the other in the lumbar curve of their back, creating three contact points with the body. The top hand is the opposite side of the body to the front foot that is placed on the board. The participant begins the test by lunging downwards until their back knee touches the board behind the heel of their front foot, ensuring that the whole front foot maintains contact with the board and the dowel maintains its three points of contact with the body. The participant completes the test by returning to the start position. The test is repeated three times on each leg.

Shoulder Mobility

The shoulder mobility subtest assesses an individual's bilateral shoulder range of motion. Participants use both shoulders, placing one in internal rotation with adduction and extension and the other in external rotation, with abduction and flexion. Prior to completing the subtest, the individuals hand length is measured from their distal crease to the tip of their third digit. To begin the test, the participant must stand with their feet together and head facing forward and body standing tall. The participant is asked to make a fist with each hand, placing their thumb inside the fist. They are then asked to maximally adduct, extend and externally rotate with one shoulder, and maximally abduct, flex and internally rotate with the other shoulder. The hands must remain in a fist and the feet together, with the head still facing forward. The participant holds the position where both shoulders are maximally rotated, and the tester measures the gap between the two closest bony prominences. The shoulder mobility subtest can be repeated up to three times on each shoulder. Upon completion of the subtest, there is a clearing test to observe if there is a pain response. If a pain is felt during this tests a score of zero is given, no matter how good the performance on the shoulder mobility subtest.

APPENDIX F.3 – DESCRIPTION OF ACTIVE STRAIGHT LEG RAISE & TRUNK STABILITY

Active Straight Leg Raise

This subtest is designed to assess hip mobility in the passive leg, whilst measuring hamstring and gastric-soleus flexibility in the active leg. Similarly, to the shoulder mobility subtest, the scoring criteria is based on a measurement linked to the participant. The participant starts in the supine position, with arms by their side and palms facing upwards. The tester locates their anterior superior iliac spine and the mid-point of their patella and measures the distance between the two, to identify the mid-thigh point. The participant is instructed to lift their test leg, keeping their ankle dorsiflexed and knee fully extended. The passive leg should remain completely still, with toes pointed upwards and head on the floor. The participant lifts their leg until they reach their end range, and at this point the position of the malleolus is noted.

Trunk Stability

The trunk stability subtest is designed to test an individual’s ability to stabilise the spine in an anterior and posterior plane during a close-chain upper body movement. The subtest has two different starting positions for the hands depending on performance, and has different starting positions for female participants. Individuals assume a prone position, with their feet together and their hands shoulder width apart, and placed in the appropriate start position. The knees are then fully extended to keep the thighs off the floor, and the ankles dorsiflexed. From this position, the participant is asked to complete one push up, without any lag in the lumbar spine, so the body is moved as one complete unit. If the individual is unable to successfully perform the push up, they resume the start position, but lower their hands to the appropriate second starting position before performing a push up once again. Upon completion of the subtest, a clearing test is completed. If pain is reported during the clearing test, then a score of zero is given.

APPENDIX F.4 – DESCRIPTION OF ROTARY STABILITY

Rotary Stability

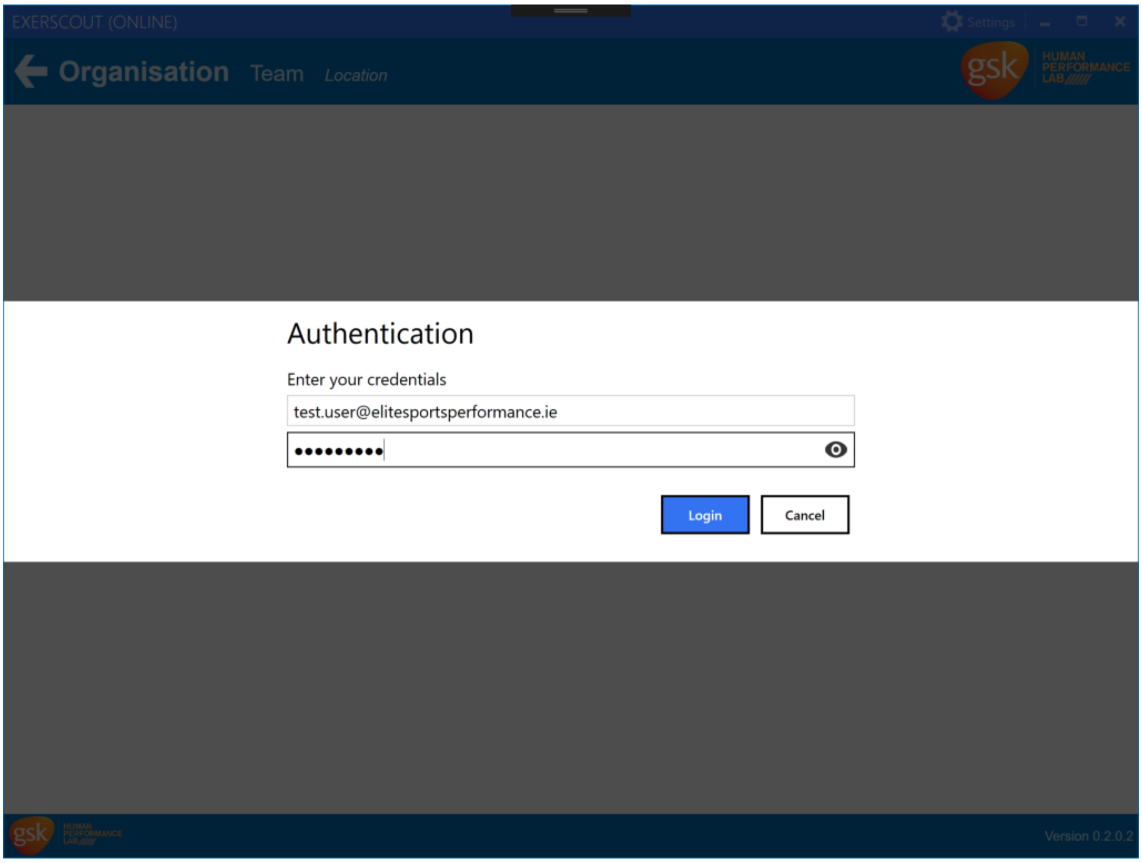
This subtest assesses multi-plane trunk stability during a combined upper and lower extremity movement. It is essentially testing an individual's ability to transfer energy from one part of the body to the other through the trunk in a co-ordinated and controlled manner. The participant begins the subtest in a quadruped position, with the FMS board set between their knees and hands and their shoulders and hips at 90° to the hands and knees respectively. The knees are set at 90° and ankles set in a dorsiflexed position. The hands and knees should remain in contact with the FMS board. The test performance differs for a score of three and a score of two. To score a three, the participant must flex one shoulder and extend the same side hip and knee. The leg and hand are only required to be raised approximately 6 inches from the floor, but the elbow, hand and knee that are lifted should remain in line with the board, along with the torso. The same shoulder is then extended, and the knee flexed so the elbow and knee touch. This movement is performed up to three times on both sides. If a score of three is not achieved, the participant performs a diagonal pattern using the opposite shoulder and hip, but in the same manner as described above. Inability to correctly perform this pattern results in a score of one being attained. A clearing test is performed after the subtest has been completed and if pain is reported during this clearing test, a score of zero is provided for test performance.

APPENDIX G

PROTOTYPE V3 DATABASE DEVELOPMENT

APPENDIX G.1 – LOGIN PAGE

As one of the objectives of the development of software was to automatically assess the FMS in an applied setting, a database was required to add functionality relating to participant and session information. Therefore, the researcher and development team worked together to establish the key information that would be needed from an applied and research perspective moving forward. To protect any personal or sensitive data, a password protected login page was created to ensure only users who should be gaining access to the software were able to log in (Figure G.1).



Authentication

Enter your credentials

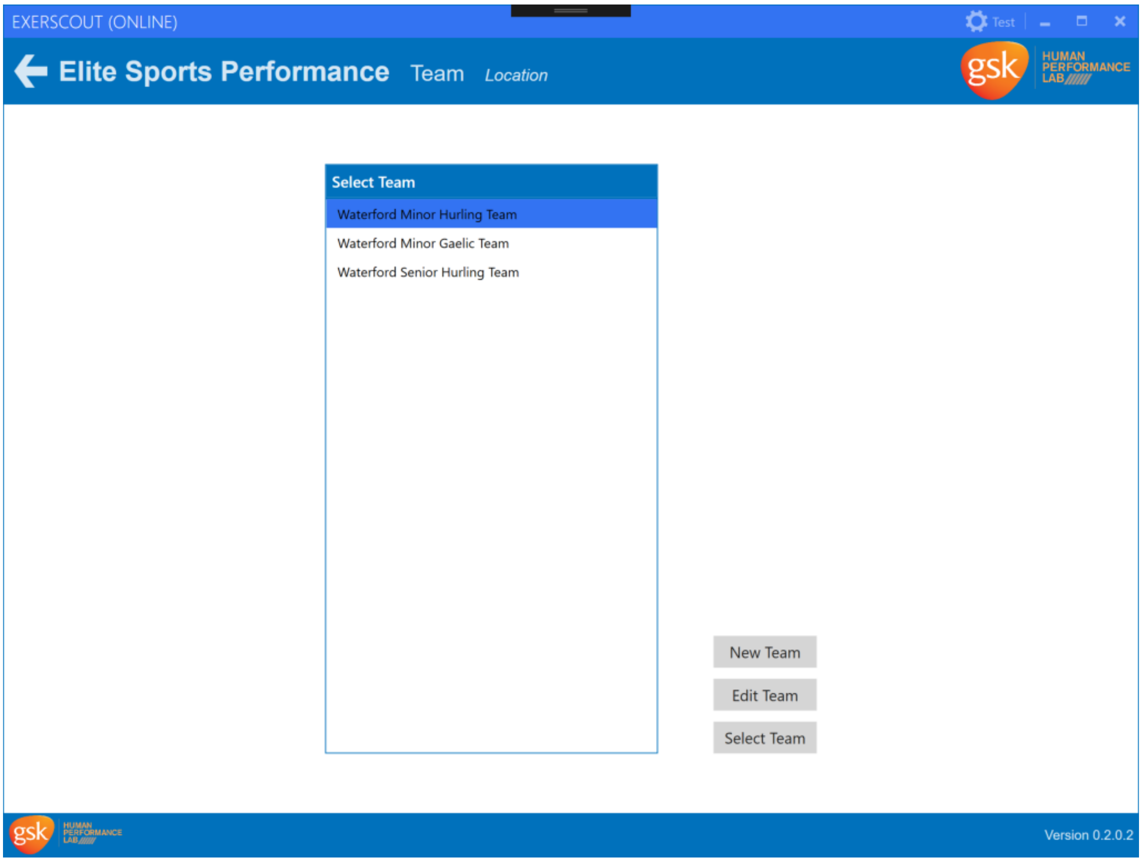
test.user@elitesportsperformance.ie

●●●●●●●●●

Figure G.1. Screenshot of login page for prototype v3

APPENDIX G.2 – TEAM SET UP PAGE

Once logged in, the user of the system needed the ability to select or set up different teams within the software depending on their own individual requirements. Therefore, a team page was created whereby users could select existing teams already stored on the system or create new teams as and when required (Figure G.2).



Waterford Minor Hurling Team
Waterford Minor Gaelic Team
Waterford Senior Hurling Team

Figure G.2. Screenshot of team page for prototype v3

APPENDIX G.3 – SESSION PAGE

Once the user had selected the team they will be testing they needed the ability to create different testing sessions for each team. This allowed them to track an individual’s progress by comparing recent performances with past performances from previous sessions (Figure G.3). The user had the option to use an existing session, thereby having the ability to add individuals who may have missed a certain session due to injury or illness, and to create new sessions at the touch of a button All sessions were date stamped and provided the option to add details relating to venue etc.

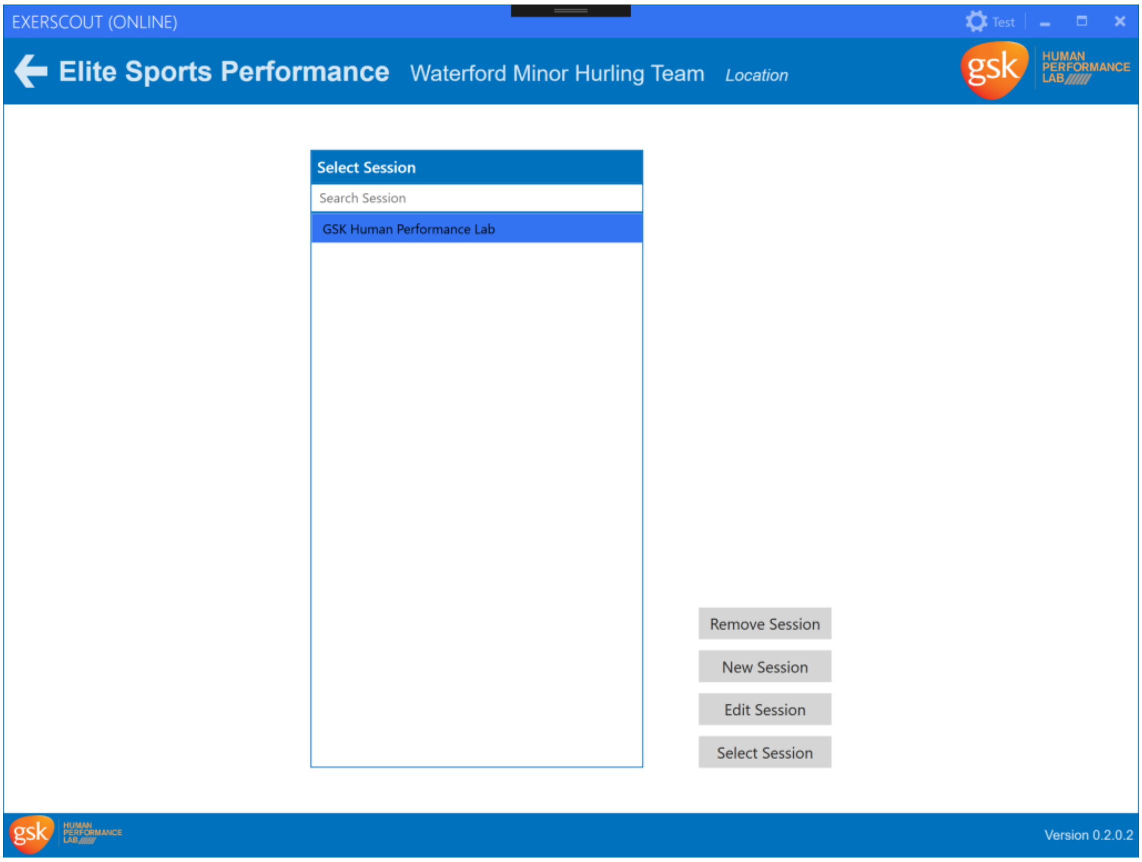


Figure G.3. Screenshot of session page for prototype v3

APPENDIX G.4 – MAIN MENU PAGE

As the number of subtests increased for Prototype v3, and the functionality was expanded to allow storage of data and results, a main menu page was designed to allow users to easily navigate to their chosen area within the software (Figure G.4).

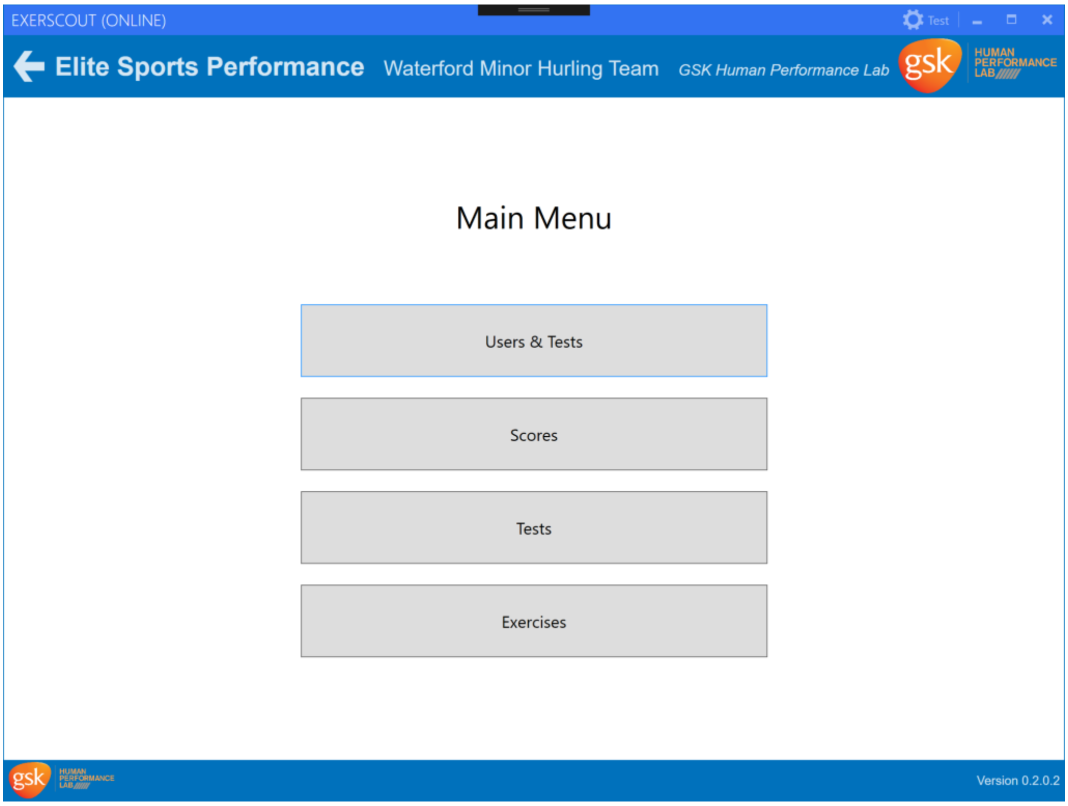


Figure G.4. Screenshot of main menu page for prototype v3

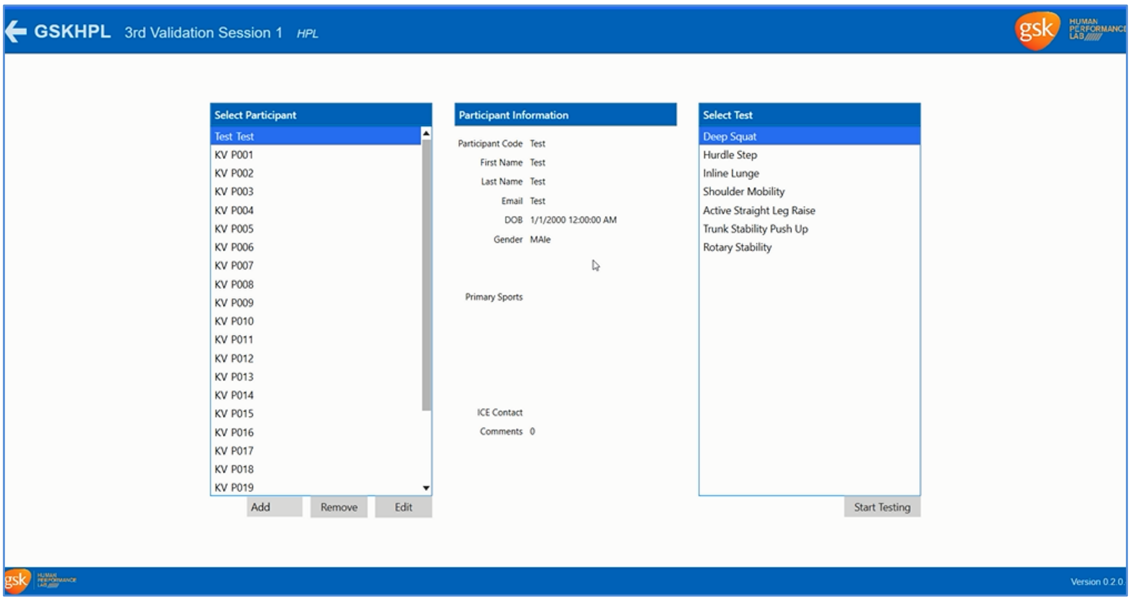
The main menu page contained four options.

- Users and Tests – Allowed user to run a testing session
- Scores – Allowed user to check scores of individuals from previous sessions
- Tests – Provided videos and instructions for subtest set up
- Exercises – Provided corrective exercises for each subtest

For the purposes of prototype v3 the “Users and Tests” and “Scores” pages were the only two pages that were operational. The other options were not available but were designed to be ready for prototype v4.

APPENDIX G.5 – USERS & TESTS PAGE

This page provided details relating to participants that had been assigned to the team selected, and allowed a user to add new participants, and edit or remove existing participants. It also allowed them to add various information relating to each participant including details such as gender, age, playing position etc (Figure G.5).

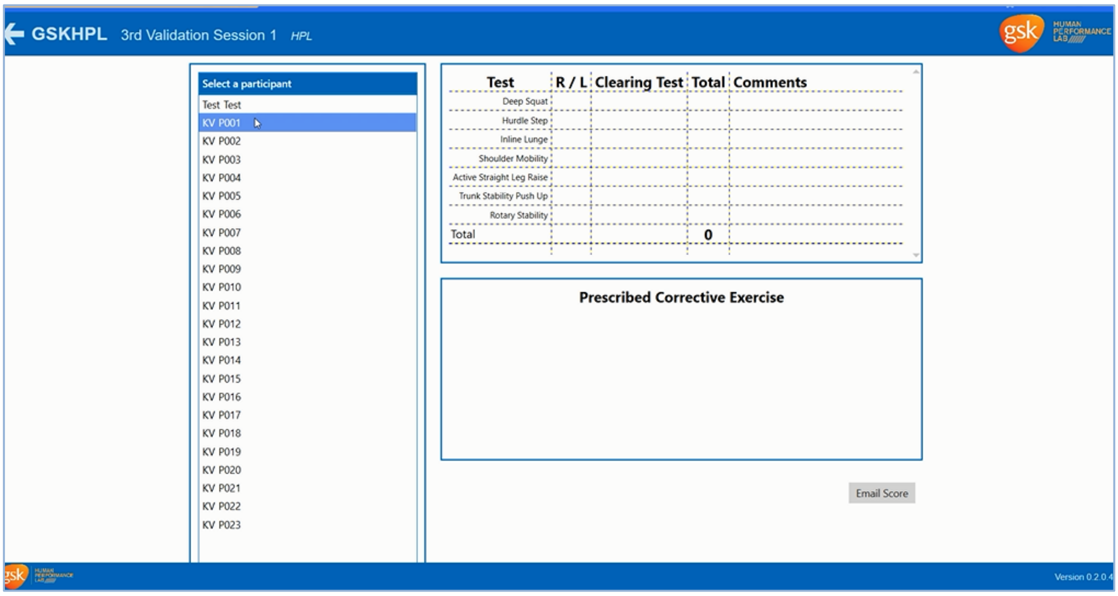


Participant Code Test
First Name Test
Last Name Test
Email Test
DOB 1/1/2000 12:00:00 AM
Gender MAle
Primary Sports
ICE Contact
Comments 0

Figure G.5. Screenshot of users and tests page for prototype v3

APPENDIX G.6 – SCORES PAGE

This page provided a breakdown of the scores achieved on each of the subtests by a participant and their total FMS composite score. The scores shown on this page were automatically saved by the software following each subtest performance and then stored in the “Scores” page (Figure G.6). As each performance on the FMS is linked to each individual who is linked to a team and session as outlined in previous paragraphs, the scores for several performances by one participant could be saved by the software.



Test	R / L	Clearing Test	Total	Comments
Deep Squat				
Hurdle Step				
Inline Lunge				
Shoulder Mobility				
Active Straight Leg Raise				
Trunk Stability Push Up				
Rotary Stability				
Total			0	

Figure G.6. Screenshot of scores page for prototype v3

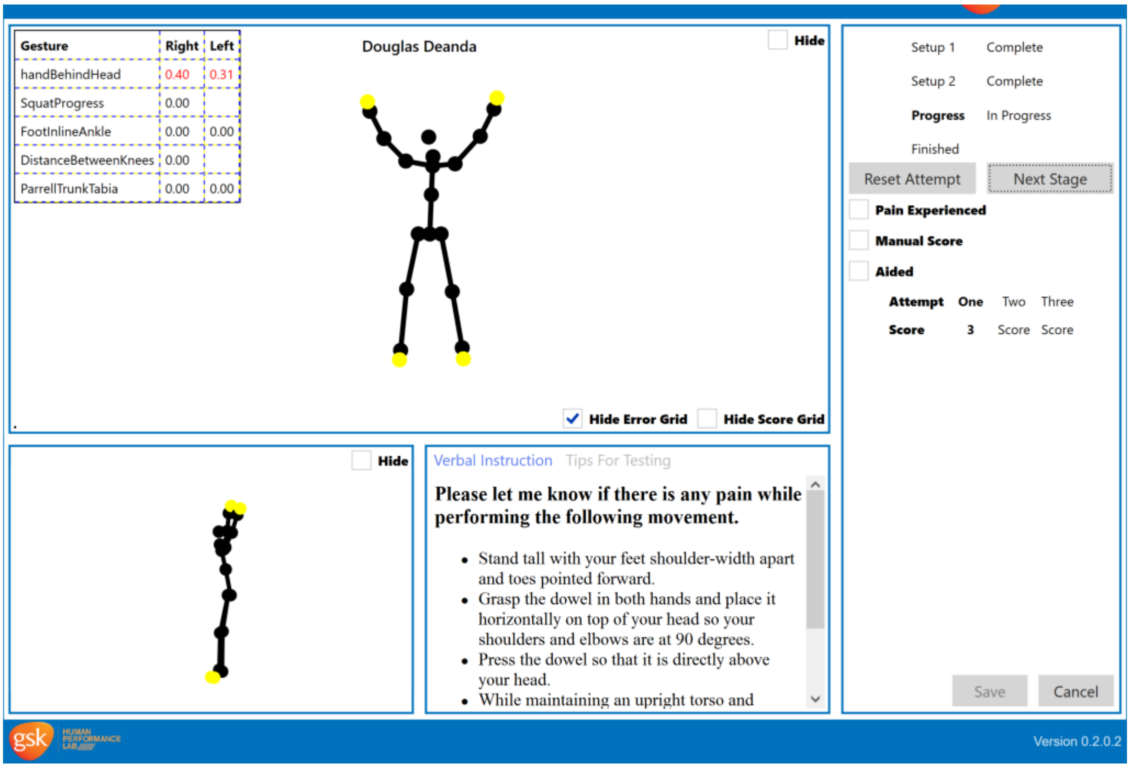
APPENDIX H

INTERFACE OPERATION PROTOTYPE

V3 & V4

APPENDIX H.1 – DEEP SQUAT

The operation of the deep squat subtest remained as per prototype v2. However, the user interface from prototype v2 was updated to enhance the user experience and provide them with the necessary level of control to correctly conduct the test in a consistent manner.



Gesture	Right	Left
handBehindHead	0.40	0.31
SquatProgress	0.00	
FootInlineAnkle	0.00	0.00
DistanceBetweenKnees	0.00	
ParrellTrunkTabia	0.00	0.00

Douglas Deanda

☐ Hide

☑ Hide Error Grid ☐ Hide Score Grid

Setup 1 Complete

Setup 2 Complete

Progress In Progress

Finished

☐ **Pain Experienced**

☐ **Manual Score**

☐ **Aided**

Attempt	**One**	Two	Three
Score	**3**	Score	Score

Verbal Instruction Tips For Testing

Please let me know if there is any pain while performing the following movement.

- Stand tall with your feet shoulder-width apart and toes pointed forward.
- Grasp the dowel in both hands and place it horizontally on top of your head so your shoulders and elbows are at 90 degrees.
- Press the dowel so that it is directly above your head.
- While maintaining an upright torso and

Figure H.1. Screenshot of deep squat subtest for prototype v3

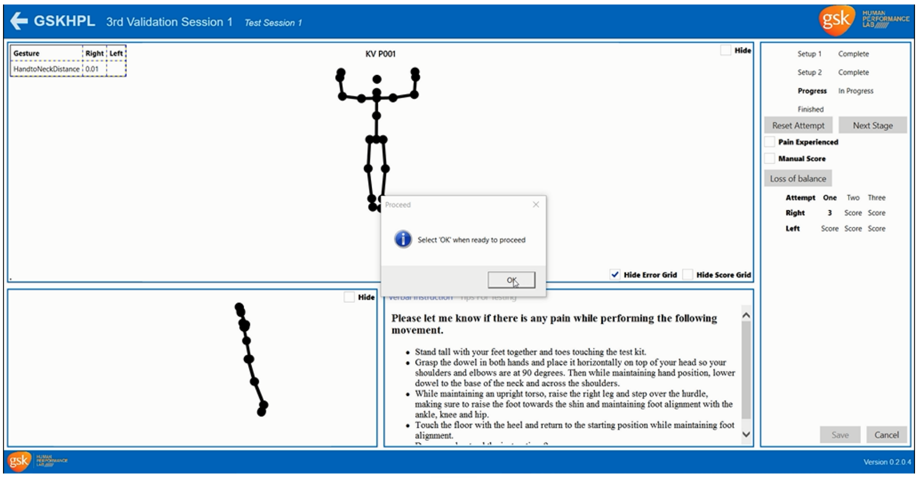
The layout of the interface was updated from that seen in prototype v2 (Figure H.1). Verbal instructions were also added to allow the user to provide the correct verbal cues to instruct a participant in how to set up correctly prior to subtest performance. The option to manually record if pain was felt by the participant was added, and if checked automatically set the score to zero for that performance. As per prototype v2, the option to record if the FMS board was in use, remained, and if checked meant the maximum score that could be achieved for the upcoming performance was a two. These updates resulted in a much simpler user interface that was easier to use and would result in more consistent and reliable testing.

APPENDIX H.2 – HURDLE STEP

A new page was designed for the hurdle step subtest, with the same user interface as the deep squat, but with slight variations included in the operation of the software and the manual elements included within the scoring system (Figure H.2). As per the deep squat, the hurdle step included a set up phase that was assessed by the software to ensure correct hand positioning prior to test performance. Upon this correct position being adopted, a pop up window appeared on screen to prompt the user to begin the test. Once the "OK" button in the pop up window was checked the software automatically began assessment of performance using the custom gestures included within the "Test" phase. Assessment continued until the user clicks the "Next Stage" button, which indicated performance was complete and assessment was halted at this point.

Before the final score for each performance was set, the user could manually indicate if the participant was in pain during performance by checking the "Pain" box. This immediately changed the scores for this performance to a zero. Once the user was confident that the performance has been fully assessed, clicking on the "Next Stage" button would save the score and move assessment on to the next performance. Following completion of all three performances on both the right and left sides, the software saved the best score from the right and left side. These scores were saved for this participant and the software used the standard FMS scoring guidelines to select the lower of the best scores provided for each side. This overall score was the score that was used to compute the total composite score for the participant, which was available to view when the user clicked the "Score" button in the main menu.

APPENDIX H.2 (CONT) – HURDLE STEP



GSKHPL 3rd Validation Session 1 *Test Session 1*

Gesture	Right	Left
HandtoNeckDistance	0.01	

KV P001

Proceed

Select 'OK' when ready to proceed

☑ Hide Error Grid ☐ Hide Score Grid

Setup 1 Complete
Setup 2 Complete
Progress In Progress
Finished

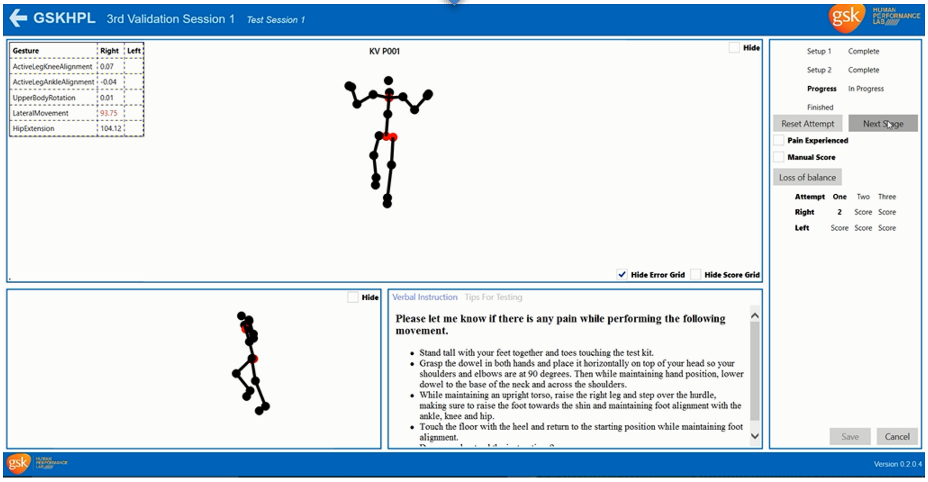
☐ **Pain Experienced**
☐ **Manual Score**

Attempt	**One**	Two	Three
Right	3	Score	Score
Left	Score	Score	Score

Please let me know if there is any pain while performing the following movement.

- Stand tall with your feet together and toes touching the test kit.
- Grasp the dowel in both hands and place it horizontally on top of your head so your shoulders and elbows are at 90 degrees. Then while maintaining hand position, lower dowel to the base of the neck and across the shoulders.
- While maintaining an upright torso, raise the right leg and step over the hurdle, making sure to raise the foot towards the shin and maintaining foot alignment with the ankle, knee and hip.
- Touch the floor with the heel and return to the starting position while maintaining foot alignment.

Version 0.2.0.4



GSKHPL 3rd Validation Session 1 *Test Session 1*

Gesture	Right	Left
ActiveLegKneeAlignment	0.07	
ActiveLegAnkleAlignment	-0.04	
UpperBodyRotation	0.01	
LateralMovement	93.75	
HipExtension	104.12	

KV P001

☑ Hide Error Grid ☐ Hide Score Grid

Setup 1 Complete
Setup 2 Complete
Progress In Progress
Finished

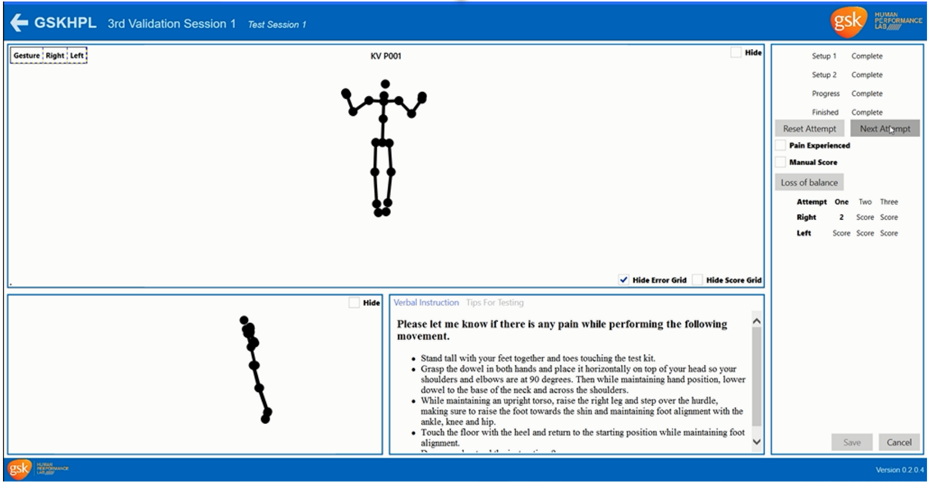
☐ **Pain Experienced**
☐ **Manual Score**

Attempt	**One**	Two	Three
Right	2	Score	Score
Left	Score	Score	Score

Please let me know if there is any pain while performing the following movement.

- Stand tall with your feet together and toes touching the test kit.
- Grasp the dowel in both hands and place it horizontally on top of your head so your shoulders and elbows are at 90 degrees. Then while maintaining hand position, lower dowel to the base of the neck and across the shoulders.
- While maintaining an upright torso, raise the right leg and step over the hurdle, making sure to raise the foot towards the shin and maintaining foot alignment with the ankle, knee and hip.
- Touch the floor with the heel and return to the starting position while maintaining foot alignment.

Version 0.2.0.4



GSKHPL 3rd Validation Session 1 *Test Session 1*

Gesture	Right	Left

KV P001

☑ Hide Error Grid ☐ Hide Score Grid

Setup 1 Complete
Setup 2 Complete
Progress Complete
Finished Complete

☐ **Pain Experienced**
☐ **Manual Score**

Attempt	**One**	Two	Three
Right	2	Score	Score
Left	Score	Score	Score

Please let me know if there is any pain while performing the following movement.

- Stand tall with your feet together and toes touching the test kit.
- Grasp the dowel in both hands and place it horizontally on top of your head so your shoulders and elbows are at 90 degrees. Then while maintaining hand position, lower dowel to the base of the neck and across the shoulders.
- While maintaining an upright torso, raise the right leg and step over the hurdle, making sure to raise the foot towards the shin and maintaining foot alignment with the ankle, knee and hip.
- Touch the floor with the heel and return to the starting position while maintaining foot alignment.

Version 0.2.0.4

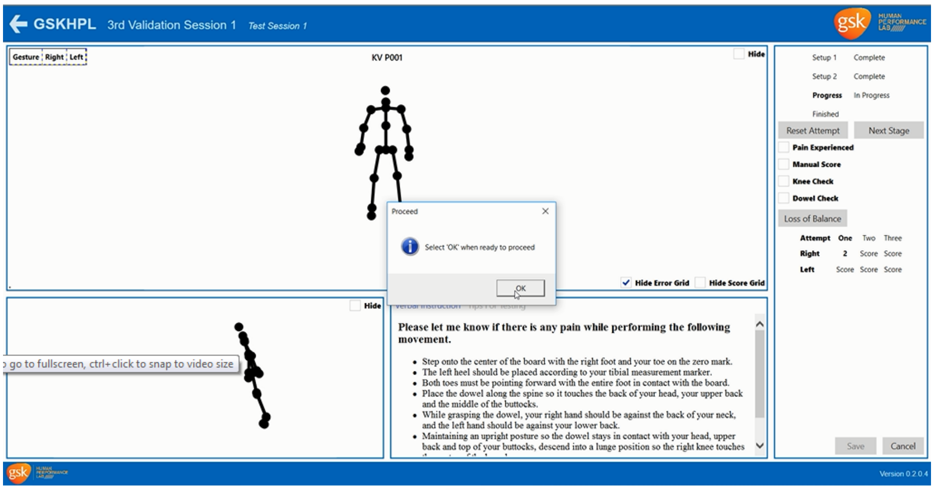
Figure H.2. Prototype v3 screenshots highlighting the process for recording hurdle step performance

APPENDIX H.3 – INLINE LUNGE

The inline lunge involved a relatively difficult set up process which was not possible to assess using prototype v3. Therefore, when a user clicked on the inline lunge subtest in the “Users and Tests” page, the software automatically brought up the pop up window to indicate it was ready to begin test assessment once the “OK” button was clicked (Figure H.3). This ensured that the participant had all the time they need to get set up ready to undertake performance, and that the software user could assist them if required. Prior to getting set up, the participant was asked to stand facing the Kinect sensor, so all the joints of the body could be correctly identified by the sensor and tracked during performance. This protocol was completed because the participant would stand on the FMS board which was set at a 30° compared to the sensor and as such increased the chances of the sensor being able to track performance. Once the “OK” button was clicked, the software automatically assessed performance using the custom gestures included in the “Test” phase. Upon test completion, the user clicked “Next Stage” and the software stopped assessment of the performance and provided a score.

Before the final score for each performance was set, the user could manually indicate if the participant was in pain during performance by checking the “Pain” box. This immediately changed the scores for this performance to a zero. In addition, the user could indicate if the participant’s knee did not touch the board during performance or if the dowel did not remain in contact with the body during performance. If either of these boxes were checked, the software automatically reduced a score of three to a two, but did not change an existing score of two as these faults did not automatically indicate a performance should be scored as a one. Once the user was confident that the performance has been fully assessed, they clicked the “Next Stage” button to save the score for that performance and moved assessment on to the next test. As per the hurdle step test, three tests were conducted on each side and the best score for each side was saved by the software for each participant. The lowest of these scores was then used to make up the total composite score for each individual.

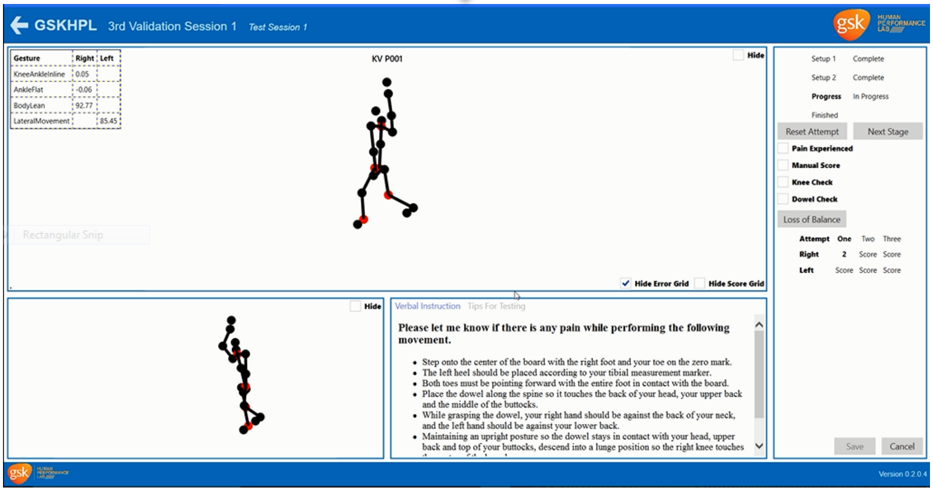
APPENDIX H.3 (CONT) – INLINE LUNGE



Select 'OK' when ready to proceed

Please let me know if there is any pain while performing the following movement.

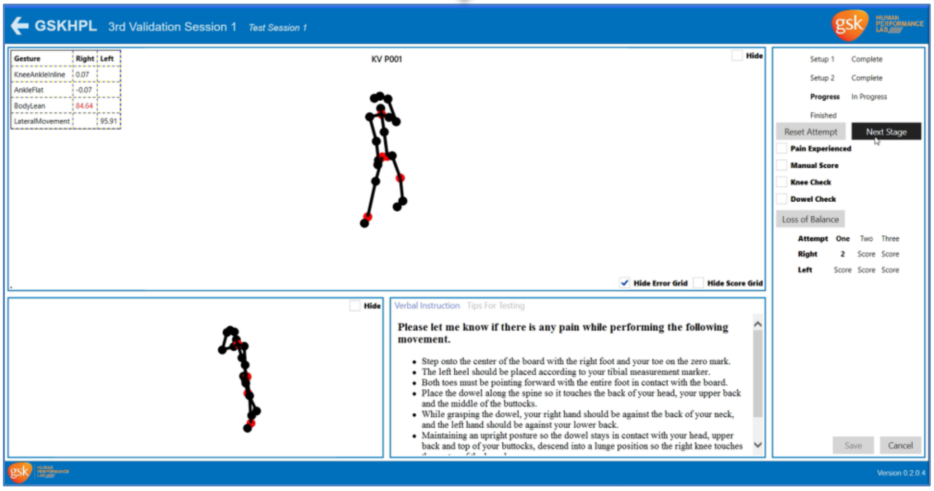
- Step onto the center of the board with the right foot and your toe on the zero mark.
- The left heel should be placed according to your tibial measurement marker.
- Both toes must be pointing forward with the entire foot in contact with the board.
- Place the dowel along the spine so it touches the back of your head, your upper back and the middle of the buttocks.
- While grasping the dowel, your right hand should be against the back of your neck, and the left hand should be against your lower back.
- Maintaining an upright posture so the dowel stays in contact with your head, upper back and top of your buttocks, descend into a lunge position so the right knee touches



Gesture	Right	Left
KneeAnkleInline	0.05	
AnkleFlat	-0.06	
BodyLean	92.77	
LateralMovement		85.45

Please let me know if there is any pain while performing the following movement.

- Step onto the center of the board with the right foot and your toe on the zero mark.
- The left heel should be placed according to your tibial measurement marker.
- Both toes must be pointing forward with the entire foot in contact with the board.
- Place the dowel along the spine so it touches the back of your head, your upper back and the middle of the buttocks.
- While grasping the dowel, your right hand should be against the back of your neck, and the left hand should be against your lower back.
- Maintaining an upright posture so the dowel stays in contact with your head, upper back and top of your buttocks, descend into a lunge position so the right knee touches



Gesture	Right	Left
KneeAnkleInline	0.07	
AnkleFlat	-0.07	
BodyLean	84.64	
LateralMovement		95.91

Please let me know if there is any pain while performing the following movement.

- Step onto the center of the board with the right foot and your toe on the zero mark.
- The left heel should be placed according to your tibial measurement marker.
- Both toes must be pointing forward with the entire foot in contact with the board.
- Place the dowel along the spine so it touches the back of your head, your upper back and the middle of the buttocks.
- While grasping the dowel, your right hand should be against the back of your neck, and the left hand should be against your lower back.
- Maintaining an upright posture so the dowel stays in contact with your head, upper back and top of your buttocks, descend into a lunge position so the right knee touches

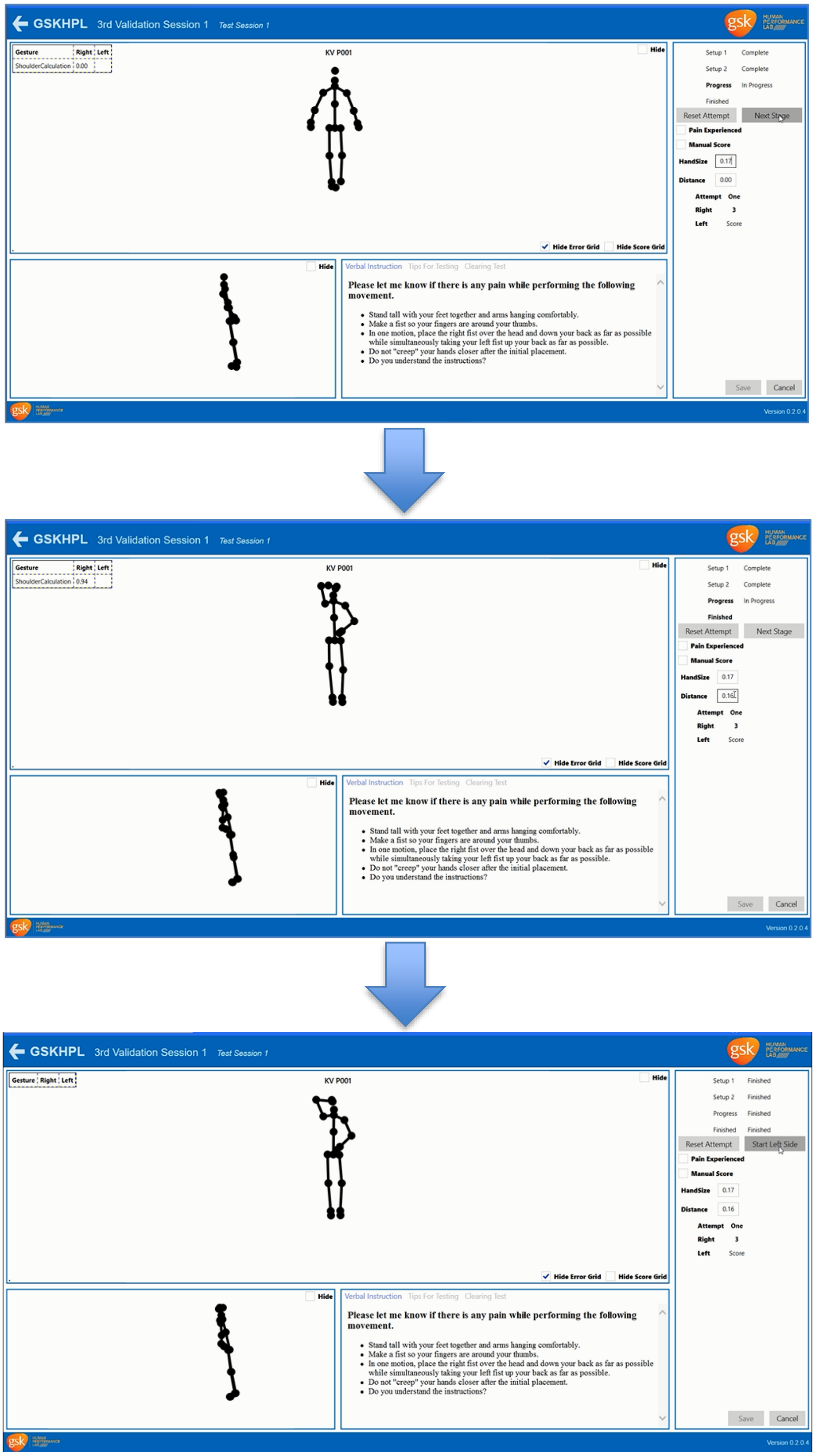
Figure H.3. Prototype v3 screenshots highlighting the process for recording inline lunge performance

APPENDIX H.4 – SHOULDER MOBILITY

The shoulder mobility subtest was set up as a semi-automated process and as such did not use the Kinect sensors “Skeletal Tracking” functionality to score the subtest. Therefore, it was not necessary for the participant to stand in view of the sensor to complete test performance. Once the tester had clicked the shoulder mobility button in the “Users and Tests” interface the shoulder mobility subtests page appeared. Unlike the other subtests, the tester was required to manually progress the test by pressing the “Next Stage” button when appropriate. Before reaching this stage, the tester is required to input the length of the participant’s hand in the “HandSize” box. This provided the software with the necessary measurement required to calculate a score once the subtest had been performed. Once the hand length had been inputted the tester could click the “Next Stage” button twice to take the subtest to the “Progress stage”. The participant was then asked to complete their performance and the tester needed to manually record the distance between the fists. This measurement was inputted in the “Distance” box in the user interface. The software was then able to automatically calculate whether the distance inputted was within one hand or one and a half hand lengths and provided the correct score accordingly.

Before the final score for each performance was set, the tester could manually indicate if the participant was in pain during performance by checking the “Pain” box. This immediately changed the scores for this performance to a zero. Once the user was confident that the performance had been fully assessed, clicking on the “Next Stage” button saved the score and moved assessment on to the next performance. Following completion of all performances on both the right and left sides, the software saved the best score from the right and left side. These scores were saved for this participant and the software used the standard FMS scoring guidelines and selected the lower of the best scores provided for each side. Prior to exiting the shoulder mobility interface, a pop up window appeared that reminded the tester to complete a clearing test and asked whether pain was present during performance of same. If the tester clicked “Yes” the software automatically recorded this on the participants final score sheet, with a zero given for shoulder mobility performance.

APPENDIX H.4 (CONT) – SHOULDER MOBILITY



Gesture	Right	Left
ShoulderCalculation	0.00	

Setup 1 Complete
Setup 2 Complete
Progress In Progress
Finished
☐ Pain Experienced
☐ Manual Score
HandSize 0.17
Distance 0.00
Attempt One
Right 3
Left Score

Please let me know if there is any pain while performing the following movement.

- Stand tall with your feet together and arms hanging comfortably.
- Make a fist so your fingers are around your thumbs.
- In one motion, place the right fist over the head and down your back as far as possible while simultaneously taking your left fist up your back as far as possible.
- Do not "creep" your hands closer after the initial placement.
- Do you understand the instructions?

Gesture	Right	Left
ShoulderCalculation	0.94	

Setup 1 Complete
Setup 2 Complete
Progress In Progress
Finished
☐ Pain Experienced
☐ Manual Score
HandSize 0.17
Distance 0.16
Attempt One
Right 3
Left Score

Please let me know if there is any pain while performing the following movement.

- Stand tall with your feet together and arms hanging comfortably.
- Make a fist so your fingers are around your thumbs.
- In one motion, place the right fist over the head and down your back as far as possible while simultaneously taking your left fist up your back as far as possible.
- Do not "creep" your hands closer after the initial placement.
- Do you understand the instructions?

Gesture	Right	Left

Setup 1 Finished
Setup 2 Finished
Progress Finished
Finished Finished
☐ Pain Experienced
☐ Manual Score
HandSize 0.17
Distance 0.16
Attempt One
Right 3
Left Score

Please let me know if there is any pain while performing the following movement.

- Stand tall with your feet together and arms hanging comfortably.
- Make a fist so your fingers are around your thumbs.
- In one motion, place the right fist over the head and down your back as far as possible while simultaneously taking your left fist up your back as far as possible.
- Do not "creep" your hands closer after the initial placement.
- Do you understand the instructions?

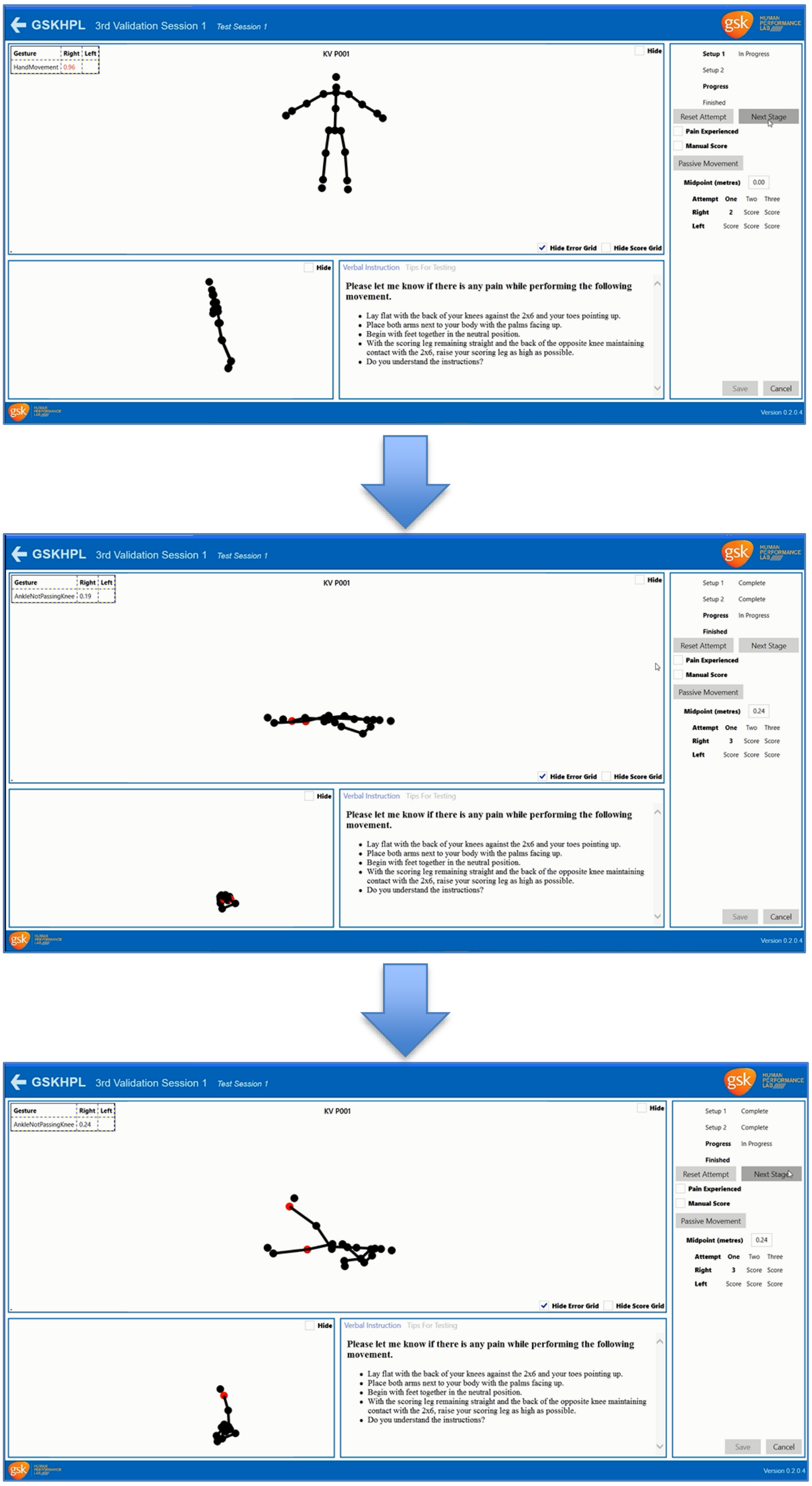
Figure H.4. Prototype v3 screenshots highlighting the process for recording shoulder mobility performance

APPENDIX H.5 – ACTIVE STRAIGHT LEG RAISE

This subtest was the first that involved the participant being placed in a supine position. Therefore, the participant was asked to stand facing the Kinect sensor in a T-Pose prior to getting into the correct set up position (Figure H.5). This allowed the sensor to identify key joint centres in the frontal plane, making it easier for it to track joint movement once the participant began to move. Prior to adopting the T-Pose, the tester was required to measure the participant's mid-thigh length. This measurement was then entered in the "Midpoint" box in the user interface, which provided the software with the information required to calculate the score once performance had taken place. Once this measurement had been inputted and the T-Pose adopted, the tester clicked the "Next Stage" button. At this point a pop up window appeared. This allowed the tester to control when the software began to assess performance once the "OK" button was pressed in the pop up window. Before beginning assessment, the participant got into the start position, which required them to be in a supine position, hands by their side and with the leg being assessed nearest the sensor. Once the "OK" button was pressed, the prototype v4 began assessment by calculating how far the active ankle moved past the passive knee. Once the participant had reached their limit, the software uses the "Midpoint" measurement to compare against the ankle position and calculated the score accordingly.

During performance, if the participant moved their passive leg or upper body to improve their score, the tester had the option to click the "Passive Movement" button. This indicated to the software that the test should end with immediate effect, and the software recorded the score reached at the point at which the button was pressed. Before the final score for each performance was set, the tester could manually indicate if the participant was in pain during performance by checking the "Pain" box. This immediately changed the score for this performance to a zero. Once the user was confident that the performance had been fully assessed, clicking on the "Next Stage" button saved the score and moved assessment on to the next performance. Following completion of all performances on both the right and left sides, the software saved the best score from the right and left side. These scores were saved for this participant and the software used the standard FMS scoring guidelines and selected the lower of the best scores provided for each side.

APPENDIX H.5 (CONT) – ACTIVE STRAIGHT LEG RAISE



Gesture	Right	Left
HandMovement	0.96	

Setup 1 In Progress
Setup 2
Progress
Finished
☐ Pain Experienced
☐ Manual Score
Midpoint (metres) 0.00

Attempt	One	Two	Three
Right	2	Score	Score
Left	Score	Score	Score

☑ Hide Error Grid ☐ Hide Score Grid

Please let me know if there is any pain while performing the following movement.

- Lay flat with the back of your knees against the 2x6 and your toes pointing up.
- Place both arms next to your body with the palms facing up.
- Begin with feet together in the neutral position.
- With the scoring leg remaining straight and the back of the opposite knee maintaining contact with the 2x6, raise your scoring leg as high as possible.
- Do you understand the instructions?

Gesture	Right	Left
AnkleNotPassingKnee	0.19	

Setup 1 Complete
Setup 2 Complete
Progress In Progress
Finished
☐ Pain Experienced
☐ Manual Score
Midpoint (metres) 0.24

Attempt	One	Two	Three
Right	3	Score	Score
Left	Score	Score	Score

☑ Hide Error Grid ☐ Hide Score Grid

Please let me know if there is any pain while performing the following movement.

- Lay flat with the back of your knees against the 2x6 and your toes pointing up.
- Place both arms next to your body with the palms facing up.
- Begin with feet together in the neutral position.
- With the scoring leg remaining straight and the back of the opposite knee maintaining contact with the 2x6, raise your scoring leg as high as possible.
- Do you understand the instructions?

Gesture	Right	Left
AnkleNotPassingKnee	0.24	

Setup 1 Complete
Setup 2 Complete
Progress In Progress
Finished
☐ Pain Experienced
☐ Manual Score
Midpoint (metres) 0.24

Attempt	One	Two	Three
Right	3	Score	Score
Left	Score	Score	Score

☑ Hide Error Grid ☐ Hide Score Grid

Please let me know if there is any pain while performing the following movement.

- Lay flat with the back of your knees against the 2x6 and your toes pointing up.
- Place both arms next to your body with the palms facing up.
- Begin with feet together in the neutral position.
- With the scoring leg remaining straight and the back of the opposite knee maintaining contact with the 2x6, raise your scoring leg as high as possible.
- Do you understand the instructions?

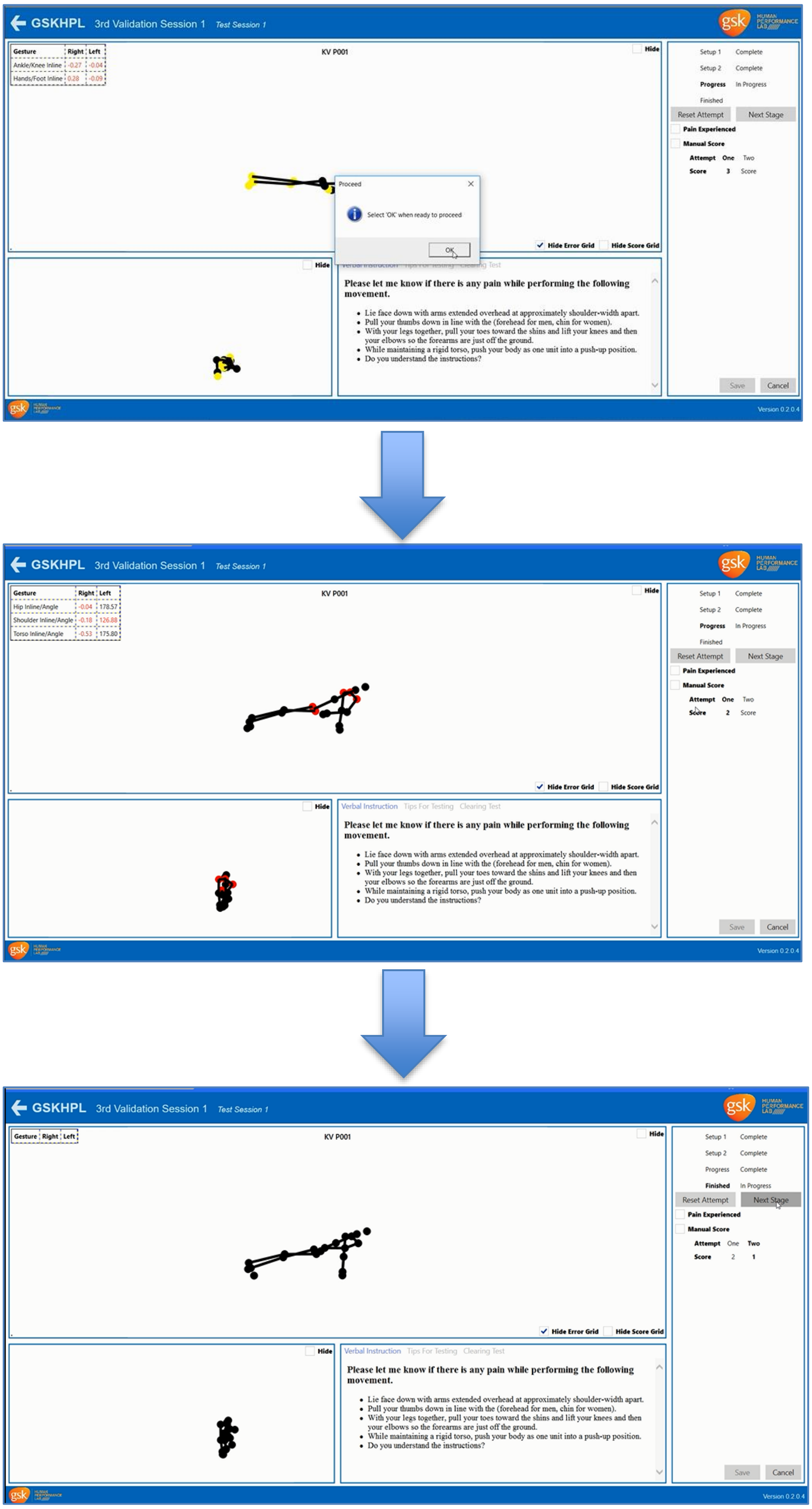
Figure H.5. Prototype v3 screenshots highlighting the process for recording active straight leg raise performance

APPENDIX H.6 – TRUNK STABILITY

The participant was asked to adopt the T-Pose position, so the sensor could identify their joints. At this point the participant assumed the start position and the hand positions were set manually by the tester. Once prototype v4 identified that they were in the correct start position, a pop up window appeared allowing the tester to control when assessment began (Figure H.6). Clicking the “OK” button began test assessment. Prototype v4 did not attempt to assess hand position or differentiate between genders performing the test. Instead it assumed that the first performance was an attempt to score a three. If the participant did not obtain a score of three on that attempt, prototype v4 assumed the next performance was an attempt to score a two. If this was not achieved, then a score of one was given. The tester had the option to restart test performance at any time, and as such had the option to allow participants three attempts at each score if required.

Before the final score for each performance was set, the tester could manually indicate if the participant was in pain during performance by checking the “Pain” box. This immediately changed the score for this performance to a zero. Once the user was confident that the performance had been fully assessed, clicking on the “Next Stage” button saved the score and moved assessment on to the next performance. The best score was saved to the overall score provided for all subtest performances. Prior to exiting the trunk stability interface, a pop up window appeared that reminded the tester to complete a clearing test and asked whether pain was present during performance of same. If the tester clicked “Yes” the software automatically recorded this on the participants final score sheet, with a zero given for trunk stability performance.

APPENDIX H.6 (CONT) – TRUNK STABILITY



Gesture	Right	Left
Ankle/Knee Inline	-0.27	-0.04
Hands/Foot Inline	0.28	-0.09

Setup 1 Complete
Setup 2 Complete
Progress In Progress
Finished

☐ Pain Experienced
☐ Manual Score

Attempt	One	Two
Score	3	Score

Proceed

Select 'OK' when ready to proceed

OK

Please let me know if there is any pain while performing the following movement.

- Lie face down with arms extended overhead at approximately shoulder-width apart.
- Pull your thumbs down in line with the (forehead for men, chin for women).
- With your legs together, pull your toes toward the shins and lift your knees and then your elbows so the forearms are just off the ground.
- While maintaining a rigid torso, push your body as one unit into a push-up position.
- Do you understand the instructions?

Gesture	Right	Left
Hip Inline/Angle	-0.04	178.57
Shoulder Inline/Angle	-0.18	126.88
Torso Inline/Angle	-0.53	175.80

Setup 1 Complete
Setup 2 Complete
Progress In Progress
Finished

☐ Pain Experienced
☐ Manual Score

Attempt	One	Two
Score	2	Score

Please let me know if there is any pain while performing the following movement.

- Lie face down with arms extended overhead at approximately shoulder-width apart.
- Pull your thumbs down in line with the (forehead for men, chin for women).
- With your legs together, pull your toes toward the shins and lift your knees and then your elbows so the forearms are just off the ground.
- While maintaining a rigid torso, push your body as one unit into a push-up position.
- Do you understand the instructions?

Gesture	Right	Left

Setup 1 Complete
Setup 2 Complete
Progress Complete
Finished In Progress

☐ Pain Experienced
☐ Manual Score

Attempt	One	Two
Score	2	1

Please let me know if there is any pain while performing the following movement.

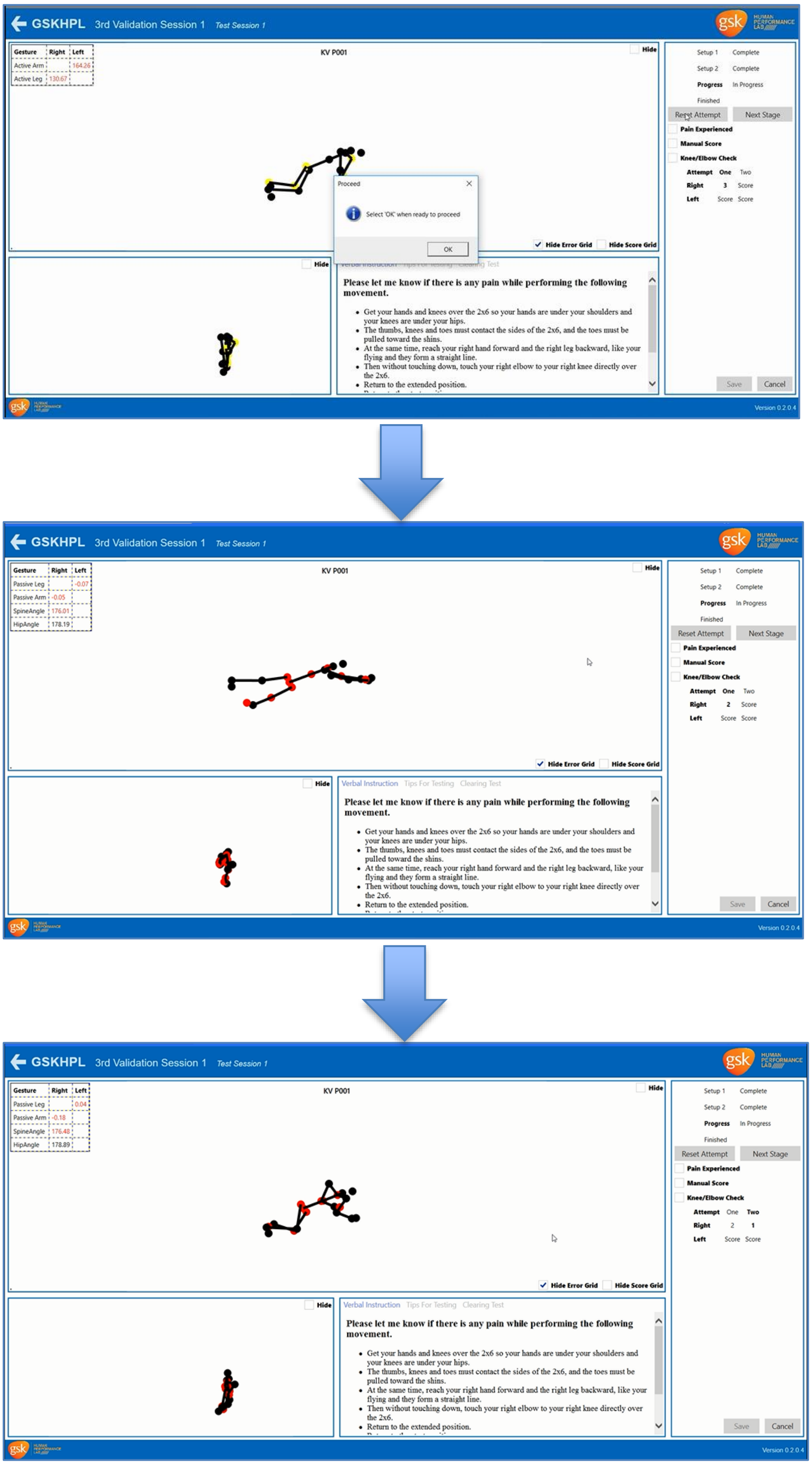
- Lie face down with arms extended overhead at approximately shoulder-width apart.
- Pull your thumbs down in line with the (forehead for men, chin for women).
- With your legs together, pull your toes toward the shins and lift your knees and then your elbows so the forearms are just off the ground.
- While maintaining a rigid torso, push your body as one unit into a push-up position.
- Do you understand the instructions?

Figure H.6. Prototype v3 screenshots highlighting the process for recording trunk stability performance

APPENDIX H.7 – ROTARY STABILITY

The final subtest was set up in the same manner as the trunk stability subtest. The participant was asked to assume a T-Pose before moving into a quadruped position where prototype v4 could identify if the limbs were correctly aligned prior to test performance. The side being assessed needed to be nearest the sensor and the tester manually checked if the knees and thumbs of the participant were touching the FMS board. Once the correct position was reached, a pop up window appeared which allowed the tester to control when assessment of performance began (Figure H.7). The software assumed that the first performance was an attempt to score a three. If this was not obtained it presumed that the second performance was an attempt to score a two, and if this was not achieved a score of one was given for that side. The tester had the option to restart each test performance, so the participant could have up to three attempts at each score if required.

During performance, the tester had the option to click the “Knee/Elbow touch” button if they observed that the participant did not touch the knee with the elbow or vice verses. If clicked the prototype reduced the score provided by one to indicate incomplete performance, no matter what score was being attempted. Before the final score for each performance was set, the tester could manually indicate if the participant was in pain during performance by checking the “Pain” box. This immediately changed the scores for that performance to a zero. Once the user was confident that the performance had been fully assessed, clicking on the “Next Stage” button saved the score and moved assessment on to the next performance. Following completion of all performances on both the right and left sides, the software saved the best score from the right and left side. These scores were saved for this participant and the software used the standard FMS scoring guidelines and selected the lower of the best scores provided for each side. Prior to exiting the rotary stability interface, a pop up window appeared that reminded the tester to complete a clearing test and asked whether pain was present during performance of same. If the tester clicked “Yes” the software automatically recorded this on the participants final score sheet, with a zero given for rotary stability performance.



Gesture	Right	Left
Active Arm		164.26
Active Leg	130.67	

KV P001

Setup 1 Complete
Setup 2 Complete
Progress In Progress
Finished

☐ Pain Experienced
☐ Manual Score
☐ Knee/Elbow Check

Attempt	One	Two
Right	3	Score
Left	Score	Score

Proceed

Select 'OK' when ready to proceed

OK

☑ Hide Error Grid ☐ Hide Score Grid

Please let me know if there is any pain while performing the following movement.

- Get your hands and knees over the 2x6 so your hands are under your shoulders and your knees are under your hips.
- The thumbs, knees and toes must contact the sides of the 2x6, and the toes must be pulled toward the shins.
- At the same time, reach your right hand forward and the right leg backward, like your flying and they form a straight line.
- Then without touching down, touch your right elbow to your right knee directly over the 2x6.
- Return to the extended position.

Gesture	Right	Left
Passive Leg		-0.07
Passive Arm	-0.05	
SpineAngle	176.01	
HipAngle	178.19	

KV P001

Setup 1 Complete
Setup 2 Complete
Progress In Progress
Finished

☐ Pain Experienced
☐ Manual Score
☐ Knee/Elbow Check

Attempt	One	Two
Right	2	Score
Left	Score	Score

☑ Hide Error Grid ☐ Hide Score Grid

Please let me know if there is any pain while performing the following movement.

- Get your hands and knees over the 2x6 so your hands are under your shoulders and your knees are under your hips.
- The thumbs, knees and toes must contact the sides of the 2x6, and the toes must be pulled toward the shins.
- At the same time, reach your right hand forward and the right leg backward, like your flying and they form a straight line.
- Then without touching down, touch your right elbow to your right knee directly over the 2x6.
- Return to the extended position.

Gesture	Right	Left
Passive Leg		0.04
Passive Arm	-0.18	
SpineAngle	176.48	
HipAngle	178.89	

KV P001

Setup 1 Complete
Setup 2 Complete
Progress In Progress
Finished

☐ Pain Experienced
☐ Manual Score
☐ Knee/Elbow Check

Attempt	One	Two
Right	2	1
Left	Score	Score

☑ Hide Error Grid ☐ Hide Score Grid

Please let me know if there is any pain while performing the following movement.

- Get your hands and knees over the 2x6 so your hands are under your shoulders and your knees are under your hips.
- The thumbs, knees and toes must contact the sides of the 2x6, and the toes must be pulled toward the shins.
- At the same time, reach your right hand forward and the right leg backward, like your flying and they form a straight line.
- Then without touching down, touch your right elbow to your right knee directly over the 2x6.
- Return to the extended position.

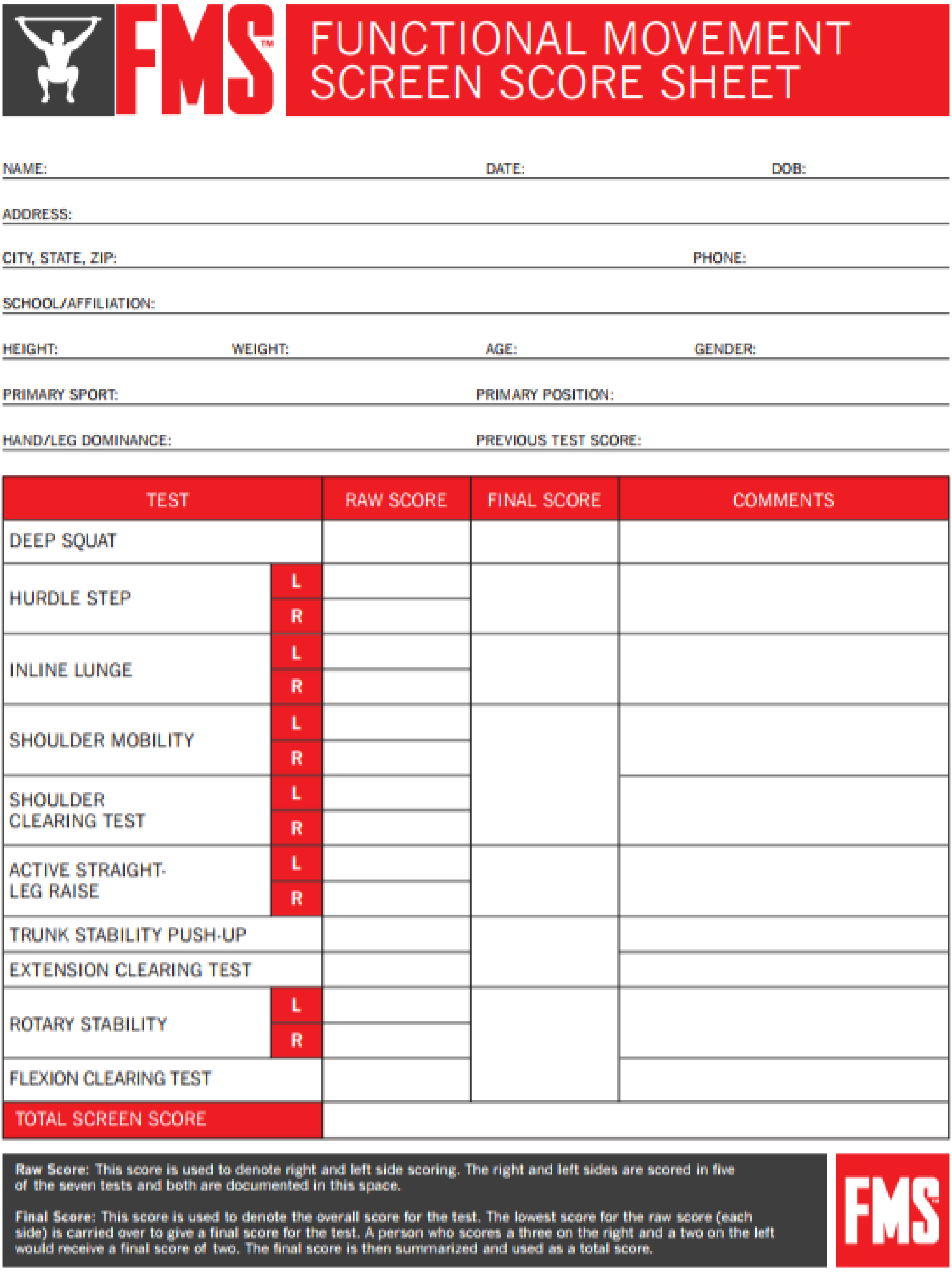
Figure H.7. Prototype v3 screenshots highlighting the process for recording rotary stability performance

APPENDIX I

STUDY THREE AND FOUR

DATA COLLECTION SHEETS

APPENDIX I.1 – FMS SCORESHEET



FMS™

FUNCTIONAL MOVEMENT SCREEN SCORE SHEET

NAME: DATE: DOB:

ADDRESS:

CITY, STATE, ZIP: PHONE:

SCHOOL/AFFILIATION:

HEIGHT: WEIGHT: AGE: GENDER:

PRIMARY SPORT: PRIMARY POSITION:

HAND/LEG DOMINANCE: PREVIOUS TEST SCORE:

TEST		RAW SCORE	FINAL SCORE	COMMENTS
DEEP SQUAT				
HURDLE STEP	L			
	R			
INLINE LUNGE	L			
	R			
SHOULDER MOBILITY	L			
	R			
SHOULDER CLEARING TEST	L			
	R			
ACTIVE STRAIGHT-LEG RAISE	L			
	R			
TRUNK STABILITY PUSH-UP				
EXTENSION CLEARING TEST				
ROTARY STABILITY	L			
	R			
FLEXION CLEARING TEST				
TOTAL SCREEN SCORE				

Raw Score: This score is used to denote right and left side scoring. The right and left sides are scored in five of the seven tests and both are documented in this space.

Final Score: This score is used to denote the overall score for the test. The lowest score for the raw score (each side) is carried over to give a final score for the test. A person who scores a three on the right and a two on the left would receive a final score of two. The final score is then summarized and used as a total score.

FMS™

APPENDIX I.2 – VIDEO SCORESHEET STUDY FOUR

Participant Code	Deep Squat	Hurdle Step Right	Hurdle Step Left	Hurdle Step Total	Inline Lunge Right	Inline Lunge Left	Inline Lunge Total	Shoulder Mobility Right	Shoulder Mobility Left	Shoulder Mobility Total	ASLR Right	ASLR Left	ASLR Total	Trunk Stability	Rotary Stability Right	Rotary Stability Left	Rotary Stability Total	Total Score	Recommended area for corrective exercise work
KVP001																		0	
KVP002																		0	
KVP003																		0	
KVP004																		0	
KVP005																		0	
KVP006																		0	
KVP007																		0	
KVP008																		0	
KVP009																		0	
KVP010																		0	
KVP011																		0	
KVP012																		0	
KVP013																		0	
KVP014																		0	
KVP015																		0	
KVP016																		0	
KVP017																		0	
KVP018																		0	
KVP019																		0	
KVP020																		0	
KVP021																		0	
KVP022																		0	
KVP023																		0	

APPENDIX I.3 – MASTER SCORESHEET STUDY FOUR

Participant Code	DS Manual	DS Software	DS Video	HS Left Manual	HS Left Software	HS Left Video	HS Right Manual	HS Right Software	HS Right Video	HS total Manual	HS Total Software	HS Total Video	IL Left Manual	IL Left Software	IL Left Video	IL Right Manual	IL Right Software	IL Right Video	IL Total Manual	IL Total Software	IL Total Video
KVP001																					
KVP002																					
KVP003																					
KVP004																					
KVP005																					
KVP006																					
KVP007																					
KVP008																					
KVP009																					
KVP010																					
KVP011																					
KVP012																					
KVP013																					
KVP014																					
KVP015																					
KVP016																					
KVP017																					
KVP018																					
KVP019																					
KVP020																					
KVP021																					
KVP022																					
KVP023																					

APPENDIX I.3 (CONT) – MASTER SCORESHEET STUDY FOUR

Participant Code	SM Left Manual	SM Left Software	SM Left Video	SM Right Manual	SM Right Software	SM Right Video	SM Total Manual	SM Total Software	SM Total Video	ASLR Left Manual	ASLR Left Software	ASLR Left Video	ASLR Right Manual	ASLR Right Software	ASLR Right Video	ASLR Total Manual	ASLR Total Software	ASLR Total Video	Trunk Stability Manual	Trunk Stability Software	Trunk Stability Video
KVP001																					
KVP002																					
KVP003																					
KVP004																					
KVP005																					
KVP006																					
KVP007																					
KVP008																					
KVP009																					
KVP010																					
KVP011																					
KVP012																					
KVP013																					
KVP014																					
KVP015																					
KVP016																					
KVP017																					
KVP018																					
KVP019																					
KVP020																					
KVP021																					
KVP022																					
KVP023																					

APPENDIX I.3 (CONT) – MASTER SCORESHEET STUDY FOUR

Participant Code	RS Left Manual	RS Left Software	RS Left Video	RS Right Manual	RS Right Software	RS Right Video	RS Total Manual	RS Total Software	RS Total Video	FMS Total Manual	FMS Total Software	FMS Total Video	Corrective Exercise Manual	Corrective Exercise Software	Corrective Exercise Video
KVP001															
KVP002															
KVP003															
KVP004															
KVP005															
KVP006															
KVP007															
KVP008															
KVP009															
KVP010															
KVP011															
KVP012															
KVP013															
KVP014															
KVP015															
KVP016															
KVP017															
KVP018															
KVP019															
KVP020															
KVP021															
KVP022															
KVP023															

APPENDIX J

PUBLICATIONS AND CONFERENCES

APPENDIX J.1 – CONFERENCES AND PUBLICATIONS

1. Smith, P. D., & Hanlon, M. P. (2015) Effectiveness of the Functional Movement Screen in predicting injury rates amongst soccer players. In: *Proceedings of the 2015 All Ireland Postgraduate Conference in Sport Sciences and Physical Education*, Limerick.

2. Smith, P. D., & Hanlon, M. P. (2015) Effectiveness of Functional Movement Screen in predicting injury rates amongst soccer players. In: *Proceedings of the 2015 Waterford Institute of Technology Research Day, Waterford.*

3. Smith, P. D., & Hanlon, M. P. (2016) Assessing the validity of using the Microsoft Kinect to automatically assess deep squat performance. In: *Proceedings of the 2016 All Ireland Postgraduate Conference in Sport Sciences and Physical Education*, *Waterford.*

4. Smith, P. D, & Hanlon, M. P. (2016) Assessing the validity of using the Microsoft Kinect v2 to automatically assess three Functional Movement Screen subtests. In: *Proceedings of the 2016 British Association of Sport and Exercise Sciences Conference, Nottingham, UK.*

5. Smith, P. D., & Hanlon, M. P. (2017) Assessing the effectiveness of the Functional Movement Screen in predicting non-contact injury rates in soccer players. In: *Journal of Strength & Conditioning Research.* 31 (12), 3327-3332.

6. Smith, P. D., & Hanlon, M. P. (2017) Validity of using the Microsoft Kinect v2 to automatically assess FMS deep squat performance. In: *Proceedings of the 2017 American College of Sports Medicine Conference, Denver, USA.*

7. Smith, P. D., & Hanlon, M. P. (2017) Assessing the validity and test-retest reliability of the Kinect sensor v2 when scoring the Functional Movement Screen. In: *Proceedings of the 2017 European College of Sports Science Conference, Essen, Germany.*

8. Smith, P. D., & Hanlon, M. P. (2018) Examining the validity of the Microsoft Kinect v2 to automatically assess three Functional Movement Screen subtests. In: *Journal of Applied Biomechanics.* Submitted for approval July 2018.

9. Smith, P. D., Flynn, M., Elwell, S., & Hanlon, M. P. (2018) Assessing the validity and test-retest reliability of the Kinect sensor v2 to assess the seven Functional Movement Screen subtests. In: *Physical Therapy in Sport.* Submitted for approval August 2018.

REFERENCES

REFERENCES

1. Aggarwal, J. K., & Cai, Q. (1999). Human Motion Analysis: A Review. *Computer Vision and Image Understanding*, *73*(3), 428–440. https://doi.org/10.1002/ps.2158

2. Alexander, E. J., & Andriacchi, T. P. (2001). Correcting for deformation in skin-based marker systems. *Journal of Biomechanics*, *34*(3), 355–361. https://doi.org/10.1016/S0021-9290(00)00192-5

3. Altun, K., Barshan, B., & Tunçel, O. (2010). Comparative study on classifying human activities with miniature inertial and magnetic sensors. *Pattern Recognition*, *43*(10), 3605–3620. https://doi.org/10.1016/j.patcog.2010.04.019

4. Appel, B. M., (2012). The Capability of the Functional Movement Screen to Predict Injury in Division I Male and Female Track and Field Athletes. *All Gradaute Plan B and Other Reports*, *Paper 174*.

5. Arnason, A., Sigurdsson, S. B., Gudmundsson, A., Holme, I., Engebretsen, L., & Bahr, R. (2004). American Journal of Sports Risk Factors for Injuries in Football. https://doi.org/10.1177/0363546503258912

6. Bacon, C. S., & Mauger, A. R. (2017). Prediction of overuse injuries in professional U18-U21 footballers using metrics of training distance and intensity. *Journal of Strength and Conditioning Research*. 31(11), 3067-3076.

7. Baltaci, G., Un, N., Tunay, V., Besler, A., & Gerçeker, S. (2003). Comparison of three different sit and reach tests for measurement of hamstring flexibility in female university students. *Br. J. Sports Med. Br J Sports Med*, *3737*, 59–6159. https://doi.org/10.1136/bjsm.37.1.59

8. Bardenett, S. M., Micca, J. J., DeNoyelles, J. T., Miller, S. D., Jenk, D. T., & Brooks, G. S. (2015). Functional Movement Screen Normative Values and Validity in High School Athletes: Can the Fms^{TM} Be Used As a Predictor of Injury? *International Journal of Sports Physical Therapy*, *10*(3), 303–8.

9. Beck, K., Thomson, J. S., Swift, R. J., & von Hurst, P. R. (2015). Role of nutrition in performance enhancement and postexercise recovery. *Open Access Journal of Sports Medicine*, 259. https://doi.org/10.2147/OAJSM.S33605

10. Bland, J. M., & Altman, D. G. (1986). Statistical methods for assessing agreement between two methods of clinical measurement. *Lancet*, *1*(8476), 307–310. https://doi.org/10.1016/S0140-6736(86)90837-8

11. Bloomfield, J., Polman, R., & O'Donoghue, P. (2007). Physical Demands of Different Positions in FA Premier League Soccer. *Journal of Sports Science & Medicine*, *6*(1), 63–70.

12. Bonnechère, B., Jansen, B., Salvia, P., Bouzahouene, H., Omelina, L., Moiseev, F., Sholukha, V., Cornelis, J., Rooze, M., & Van Sint Jan, S. (2014). Validity and reliability of the Kinect within functional assessment activities: Comparison with standard stereophotogrammetry. *Gait and Posture*, *39*(1), 593–598. https://doi.org/10.1016/j.gaitpost.2013.09.018.

13. Butler, R. J., Plisky, P. J., Southers, C., Scoma, C., & Kiesel, K. B. (2010). Biomechanical analysis of the different classifications of the Functional Movement Screen deep squat test. *Sports Biomechanics / International Society of Biomechanics in Sports*, *9*(4), 270–279. https://doi.org/10.1080/14763141.2010.539623

14. Cao, Y., Li, B. Z., Li, Q. N., Xie, J. D., Cao, B. Z., & Yu, S. Y., (2017). Kinect-based gait analyses of patients with Parkinson’s disease, patients with stroke with hemiplegia, and healthy adults. *CNS Neuroscience & Therapeutics*, (February), 1–3. https://doi.org/10.1111/cns.12687

15. Carling, C., Bloomfield, J., Nelsen, L., & Reilly, T. (2008). The Role of Motion Analysis in Elite Soccer. *Sports Medicine*. 38(10), 839–862. https://doi.org/10.2165/00007256-200838100-00004

16. Chalmers, S., Fuller, J., Debenedictis, T., Townsley, S., Lynagh, M., Gleeson, C., Zachariab, A., Thomson, S., Magarey, M., (2016). Asymmetry during preseason Functional Movement Screen testing is associated with injury during a junior Australian football season. *Journal of Science and Medicine in Sport.* https://doi.org/10.1016/j.jsams.2016.12.076

17. Chapman, R., Laymon, A. S., & Arnold, T. (2013). Functional Movement Scores and Longitudinal Performance Outcomes in Elite Track and Field Athletes. *International Journal of Sports Physiology and Performance*. 9 (2), 203-211.

18. Chorba, R. S., Chorba, D. J., Bouillon, L. E., Overmyer, C. A., & Landis, J. A. (2010). Use of a functional movement screening tool to determine injury risk in female collegiate athletes. *North American Journal of Sports Physical Therapy.* 5(2), 47–54.

19. Clark, R. A., Pua, Y. H., Fortin, K., Ritchie, C., Webster, K. E., Denehy, L., & Bryant, A. L. (2012). Validity of the Microsoft Kinect for assessment of postural control. *Gait and Posture*, *36*(3), 372–377. https://doi.org/10.1016/j.gaitpost.2012.03.033

20. Clarsen, B., Myklebust, G., & Bahr, R. (2013). Development and validation of a new method for the registration of overuse injuries in sports injury epidemiology: the Oslo Sports Trauma Research Centre (OSTRC) overuse injury questionnaire. *British Journal of Sports Medicine*, *47*(8), 495–502. https://doi.org/10.1136/bjsports-2012-091524

21. Comfort, P., Green, C. M., & Matthews, M. (2009). Training Considerations after Hamstring Injury in Athletes. *Strength & Conditioning Journal*, *31*(1).

22. Conlon, J. K. (2013). The Relationship Between the Functional Movement Screen ™ and Countermovement Jump Height. *Theses and Dissertations. Paper 213.*

23. Cook, G., Burton, L., & Hoogenboom, B. (2006). Pre-participation screening: the use of fundamental movements as an assessment of function - part 1. *North American Journal of Sports Physical Therapy.* 1(2), 62–72.

24. Cook, G., Burton, L., & Hoogenboom, B. (2006). Pre-participation screening: the use of fundamental movements as an assessment of function - part 2. *North American Journal of Sports Physical Therapy.* 1(3), 132–9.

25. Dempsey, G. M., Gibson, N. V, Sykes, D., Pryjmachuk, B., & Turner, A. P. (2017). Match demands of Senior and Junior players during International Rugby League. *Journal of Strength and Conditioning Research*, 1. https://doi.org/10.1519/JSC.0000000000002028

26. Dudley, L. A., Smith, C. A., Olson, B. K., Chimera, N. J., Schmitz, B., & Warren, M. (2013). Interrater and Intrarater Reliability of the Tuck Jump Assessment by Health Professionals of Varied Educational Backgrounds. *Journal of Sports Medicine (Hindawi Publishing Corporation)*, *2013*, 483503. https://doi.org/10.1155/2013/483503

27. Duke, S. R., Martin, S. E., & Gaul, C. A. (2017). Preseason Functional Movement Screen^{TM} Predicts Risk of Time-loss Injury in Experienced Male Rugby Union Athletes. *Journal of Strength and Conditioning Research*, 1. https://doi.org/10.1519/JSC.0000000000001838

28. Dutta, T. (2012). Evaluation of the Kinect sensor for 3-D kinematic measurement in the workplace. *Applied Ergonomics*, *43*(4), 645–649. https://doi.org/10.1016/j.apergo.2011.09.011

29. Ehrig, R. M., Taylor, W. R., Duda, G. N., & Heller, M. O. (2006). A survey of formal methods for determining the centre of rotation of ball joints. *Journal of Biomechanics*, *39*(15), 2798–2809. https://doi.org/10.1016/j.jbiomech.2005.10.002

30. Ehrig, R. M., Taylor, W. R., Duda, G. N., & Heller, M. O. (2007). A survey of formal methods for determining functional joint axes. *Journal of Biomechanics*, *40*(10), 2150–2157. https://doi.org/10.1016/j.jbiomech.2006.10.026

31. Ekstrand, J. (2008). Epidemiology of football injuries. *Science & Sports*, *23*(2), 73–77. https://doi.org/10.1016/j.scispo.2007.10.012

32. Ekstrand, J., Hägglund, M., & Waldén, M. (2009). Injury incidence and injury patterns in professional football - the UEFA Study. *British Journal of Sports Medicine.* https://doi.org/10.1136/bjsm.2009.060582

33. Fernández-Baena, A., Susín, A., & Lligadas, X. (2012). Biomechanical validation of upper-body and lower-body joint movements of kinect motion capture data for rehabilitation treatments. *Proceedings of the 2012 4th International Conference on Intelligent Networking and Collaborative Systems, INCoS 2012*, 656–661. https://doi.org/10.1109/iNCoS.2012.66

34. Freckleton, G., Cook, J., & Pizzari, T. (2014). The predictive validity of a single leg bridge test for hamstring injuries in Australian Rules Football Players. *British Journal of Sports Medicine*, *48*(8), 833.

35. Frohm, a, Heijne, a, Kowalski, J., Svensson, P., & Myklebust, G. (2012). A nine-test screening battery for athletes: a reliability study. *Scandinavian Journal of Medicine & Science in Sports*, *22*(3), 306–15. https://doi.org/10.1111/j.1600-0838.2010.01267.x

36. Frost, D. M., Beach, T. A. C., Callaghan, J. P., & McGill, S. M. (2012). Using the Functional Movement Screen^{TM} to Evaluate the Effectiveness of Training. *The Journal of Strength & Conditioning Research*, *26*(6).

37. Fuller, C. W., Ekstrand, J., Junge, A., Andersen, T. E., Bahr, R., Dvorak, J., Hagglund, P., McCrory, P., Meeuwisse, W. H. (2006). Consensus statement on injury definitions and data collection procedures in studies of football (soccer) injuries. *British Journal of Sports Medicine*, *40*(3), 193–201. https://doi.org/10.1136/bjsm.2005.025270

38. Functional Movement Screen Manual. (2017). Retrieved from https://www.functionalmovement.com/system/fms

39. Functional Movement Screen Corrective Exercises. (2017). Retrieved from https://www.functionalmovement.com/exercises

40. Gabbe, B. J., Finch, C. F., Bennell, K. L., & Wajswelner, H. (2005). Risk factors for hamstring injuries in community level Australian football. *British Journal of Sports Medicine*, *39*(2), 106–10. https://doi.org/10.1136/bjsm.2003.011197

41. Garrison, M., Westrick, R., Johnson, M. R., & Benenson, J. (2015). Association between the functional movement screen and injury development in college athletes. *International Journal of Sports Physical Therapy*, *10*(1), 21–8.

42. Glaws, K. R., Juneau, C. M., Becker, L. C., Di Stasi, S. L., & Hewett, T. E. (2014). Intra- and inter-rater reliability of the selective functional movement assessment (sfma). *International Journal of Sports Physical Therapy*, *9*(2), 195–207.

43. Gray, A. J., Jenkins, D., Andrews, M. H., Taaffe, D. R., Glover, M. L. (2010). Validity and reliability of GPS for measuring distance travelled in field based team sports. *Journal of Sports Sciences*, 28(12). https://doi.org/10.1080/02640414.2010.504783

44. Greig, M., & Siegler, J. C. (2009). Soccer-specific fatigue and eccentric hamstrings muscle strength. *Journal of Athletic Training*, *44*(2), 180–184. https://doi.org/10.4085/1062-6050-44.2.180.

45. Gulgin, H., & Hoogenboom, B. (2014). The functional movement screen (fms)^{TM}: an inter-rater reliability study between raters of varied experience. *International Journal of Sports Physical Therapy*, *9*(1), 14–20.

46. Hägglund, M., Waldén, M., Bahr, R., & Ekstrand, J. (2005). Methods for epidemiological study of injuries to professional football players: developing the UEFA model. *British Journal of Sports Medicine*, *39*(6), 340–6. https://doi.org/10.1136/bjsm.2005.018267

47. Hägglund, M., Waldén, M., & Ekstrand, J. (2006). Previous injury as a risk factor for injury in elite football: a prospective study over two consecutive seasons. *British Journal of Sports Medicine*, *40*(9), 767–72. https://doi.org/10.1136/bjsm.2006.026609

48. Hägglund, M., Waldén, M., & Ekstrand, J. (2013). Risk Factors for Lower Extremity Muscle Injury in Professional Soccer: The UEFA Injury Study, (41), 327–335.

49. Halson, S. L. (2014). Monitoring Training Load to Understand Fatigue in Athletes. *Sports Medicine,* 44(2),139–147 https://doi.org/10.1007/s40279-014-0253-z

50. Hawkins, R. D., & Fuller, C. W. (1999). A prospective epidemiological study of injuries in four English professional football clubs. *British Journal of Sports Medicine*, *33*(3), 196–203. https://doi.org/10.1136/bjsm.33.3.196.

51. Hawkins, R. D., Hulse, M., Wilkinson, C., Hodson, A., & Gibson, M. (2001). The association football medical research programme: an audit of injuries in professional football. *British Journal of Sports Medicine*, *35*(1), 43–7.

52. Heiderscheit, C, B., Sherry, M, A., Silder, A., Chumanov, E, S., Thelen, D, G,. (2010). Hamstring Strain Injuries: Recommendations for Diagnosis, Rehabilitation and Injury Prevention. *Journal of Orthopaedic Sports Physical Therapy; 40(2): 67–81. doi:10.2519/jospt.2010.3047*

53. Henderson, G., Barnes, C. A, & Portas, M. D. (2010). Factors associated with increased propensity for hamstring injury in English Premier League soccer players. *Journal of Science and Medicine in Sport / Sports Medicine Australia*, *13*(4), 397–402. https://doi.org/10.1016/j.jsams.2009.08.003

54. Herda, L., Fua, P., Plankers, R., Boulic, R., & Thalmann, D. (2001). Using skeleton-based tracking to increase the reliability of optical motion capture. *Human Movement Science*, *20*(3), 313–341. https://doi.org/10.1016/S0167-9457(01)00050-1

55. Herrington, L., & Munro, A. (2010). Drop jump landing knee valgus angle; normative data in a physically active population. *Physical Therapy in Sport : Official Journal of the Association of Chartered Physiotherapists in Sports Medicine*, *11*(2), 56–9. https://doi.org/10.1016/j.ptsp.2009.11.004

56. Herrington, L., Munro, A., & Comfort, P. (2015). A preliminary study into the effect of jumping-landing training and strength training on frontal plane projection angle. *Manual Therapy*, *20*(5), 680–685. https://doi.org/10.1016/j.math.2015.04.009

57. Herrington, L., Myer, G. D., & Munro, A. (2013). Intra and inter-tester reliability of the tuck jump assessment. *Physical Therapy in Sport*, *14*(3), 152–155. https://doi.org/10.1016/j.ptsp.2012.05.005

58. Hewett, T. E., Myer, G. D., Ford, K. R., Heidt Jr, R. S., & Colosimo, A. J. (2005). Biomechanical Measures of Neuromuscular Control and Valgus Loading of the Knee Predict Anterior Cruciate Ligament Injury Risk in Female Athletes A Prospective Study Biomechanical Measures of Neuromuscular Control and Valgus Loading of the Knee Predict Ant. *Am J Sports Med*, *33*(4), 492–501. https://doi.org/10.1177/0363546504269591

59. Higgins, T., Greene, D., & Baker, M. (2017). Effects of cold water immersion and contrast water therapy for recovery from team sport: A systematic review and meta-analysis. *Journal of Strength & Conditioning Research*, *31*(5), 1443–1460.

60. Higham, D. G., Pyne, D. B., Anson, J. M., & Eddy, A. (2012). Movement patterns in rugby sevens: Effects of tournament level, fatigue and substitute players. *Journal of Science and Medicine in Sport*, *15*(3), 277–282. https://doi.org/10.1016/j.jsams.2011.11.256

61. Jensen, U., Weilbrenner, F., Rott, F., & Eskofier, B. (2013). Sensor-based Mobile Functional Movements. *Lecture Notes of the Institute for Computer Sciences, Social Informatics and Telecommunications Engineering*, *61*(215–223).

62. Jones, P. A., Herrington, L. C., & Graham-Smith, P. (2015). Technique determinants of knee joint loads during cutting in female soccer players. *Human Movement Science*, *42*, 203–211. https://doi.org/10.1016/j.humov.2015.05.004

63. Junge, A., & Dvorak, J. (2013). Injury surveillance in the World Football Tournaments 1998-2012. *British Journal of Sports Medicine*, *47*(12), 782–8. https://doi.org/10.1136/bjsports-2013-092205

64. Junker, H., Amft, O., Lukowicz, P., & Tröster, G. (2008). Gesture spotting with body-worn inertial sensors to detect user activities. *Pattern Recognition*, *41*(6), 2010–2024. https://doi.org/10.1016/j.patcog.2007.11.016

65. Kiesel, K., Plisky, P., & Butler, R. (2011). Functional movement test scores improve following a standardized off-season intervention program in professional football players. *Scandinavian Journal of Medicine & Science in Sports*, *21*(2), 287–92. https://doi.org/10.1111/j.1600-0838.2009.01038.x

66. Kiesel, K., Plisky, P. J., & Voight, M. L. (2007). Can Serious Injury in Professional Football be Predicted by a Preseason Functional Movement Screen? *North American Journal of Sports Physical Therapy : NAJSPT*, *2*(3), 147–158.

67. Kolahi, A., Hoviattalab, M., Rezaeian, T., Alizadeh, M., Bostan, M., & Mokhtarzadeh, H. (2007). Design of a marker-based human motion tracking system, *Biomedical Signal Processing & Control, 2*(1) 59–67. https://doi.org/10.1016/j.bspc.2007.02.001

68. Leeder, J. E., Horsley, I. G., & Herrington, L. C. (2016). The Inter-rater Reliability of the Functional Movement Screen Within an Athletic Population Using Untrained Raters. *Journal of Strength and Conditioning Research / National Strength & Conditioning Association*, *30*(9), 2591–9. https://doi.org/10.1519/JSC.0b013e3182a1ff1d

69. Letafatkar, A., Hadadnezhad, M., & Shojaedin, S. (2014). Relationship between functional movement screening score and history of injury. *The International Journal of Sports Physical Therapy*, *9*(1), 21–28.

70. Lorenz, D., & Reiman, M. (2011). The role and implementation of eccentric training in athletic rehabilitation: tendinopathy, hamstring strains, and acl reconstruction. *International Journal of Sports Physical Therapy*, *6*(1), 27–44.

71. Ma, A. J., Rawat, N., Reiter, A., Shrock, C., Zhan, A., Stone, A., Rabiee, A., Griffin, S., Needham, D, M., Saria, S. (2017). Measuring Patient Mobility in the ICU Using a Novel Noninvasive Sensor. *Critical Care Medicine*, *45*(4), 630–636. https://doi.org/10.1097/CCM.0000000000002265

72. Marshall, D. A., Lopatina, E., Lacny, S., & Emery, C. A. (2016). Economic impact study: neuromuscular training reduces the burden of injuries and costs compared to standard warm-up in youth soccer. *British Journal of Sports Medicine*, 50(22) 1388-1393. https://doi.org/10.1136/bjsports-2015-095666

73. Mayorga-Vega, D., Merion-Marban, R., & Viciana, J. (2014). Criterion-related validity of sit-and-reach tests for estimating hamstring and lumbar extensibility among elderly: A Meta-Analysis. *Journal of Sports Science and Medicine,* 13, 01-14.

74. McKeown, I., Taylor-McKeown, K., Woods, C., & Ball, N. (2014). Athletic ability assessment: a movement assessment protocol for athletes. *International Journal of Sports Physical Therapy*, *9*(7), 862–73.

75. Mentiplay, B. F., Perraton, L. G., Bower, K. J., Pua, Y. H., McGaw, R., Heywood, S., & Clark, R. A. (2015). Gait assessment using the Microsoft Xbox One Kinect: Concurrent validity and inter-day reliability of spatiotemporal and kinematic variables. *Journal of Biomechanics*, 48, 2166-2170. https://doi.org/10.1016/j.jbiomech.2015.05.021

76. Minick, K., Kiesel, K., Burton, L., Taylor, A., Plisky, P., Butler, R. (2010). Interrater Reliability of the Functional Movement Screen. *Journal of Strength & Conditioning Research*, *24*(2), 479–486.

77. Moeslund, T. B., & Granum, E. (2001). A Survey of Computer Vision-Based Human Motion Capture. *Computer Vision and Image Understanding*, *81*(3), 231–268. https://doi.org/10.1006/cviu.2000.0897

78. Mohr, M., Krustrup, P., & Bangsbo, J. (2005). Fatigue in soccer: a brief review. *Journal of Sports Sciences*, *23*(6), 593–9. https://doi.org/10.1080/02640410400021286

79. Morlock, M., Windolf, M., & Gotzen, N. (2008). Systematic accuracy and precision analysis of video motion capturing systems — exemplified on the Vicon-460 system. *Journal of Biomechanics*, *41*, 2776–2780. https://doi.org/10.1016/j.jbiomech.2008.06.024

80. Muller, B., Ilg, W., Giese, M. A., & Ludolph, N. (2017). Validation of enhanced kinect sensor based motion capturing for gait assessment. *PLoS ONE*, *12*(4), 14–16.

81. Mündermann, L., Corazza, S., & Andriacchi, T. P. (2006). The evolution of methods for the capture of human movement leading to markerless motion capture for biomechanical applications. *Journal of Neuroengineering and Rehabilitation*, *3*, 6. https://doi.org/10.1186/1743-0003-3-6

82. Munro, A., Herrington, L., & Comfort, P. (2012). Comparison of landing knee valgus angle between female basketball and football athletes: Possible implications for anterior cruciate ligament and patellofemoral joint injury rates. *Physical Therapy in Sport*, *13*(4), 259–264. https://doi.org/10.1016/j.ptsp.2012.01.005

83. Myer, G. D., Ford, K. R., & Hewett, T. E. (2008). Tuck Jump Assessment for Reducing Anterior Cruciate Ligament Injury Risk. *Athletic Therapy Today*, *13*(5), 39–44. https://doi.org/10.1123/att.13.5.39

84. Noyes, F. R. (2005). The Drop-Jump Screening Test: Difference in Lower Limb Control By Gender and Effect of Neuromuscular Training in Female Athletes. *American Journal of Sports Medicine*, *33*(2), 197–207. https://doi.org/10.1177/0363546504266484

85. Nuesch, C., Roos, E., Pagenstert, G., & Mundermann, A. (2017). Measuring joint kinematics of treadmill walking and running: Comparison between an inertial sensor based system and a camera-based system. *Journal of Biomechanics*. https://doi.org/10.1016/j.jbiomech.2017.03.015

86. O’Connor, F. G., Deuster, P. A, Davis, J., Pappas, C. G., & Knapik, J. J. (2011). Functional movement screening: predicting injuries in officer candidates. *Medicine and Science in Sports and Exercise*, *43*(12), 2224–30. https://doi.org/10.1249/MSS.0b013e318223522d

87. Okada, T., Huxel, K. C., & Nesser, T. W. (2011). Relationship between core stability, functional movement and performance. *Journal of Strength & Conditioning Research*, *25*(1), 252–261. https://doi.org/10.1519/JSC.0b013e3181b22b3e

88. Parchmann, C. J., & McBride, J. M. (2011). Relationship Between Functional Movement Screen and Athletic Performance. *The Journal of Strength & Conditioning Research*, *25*(12) 3378-84.

89. Parenteau-G, E., Gaudreault, N., Chambers, S., Boisvert, C., Grenier, A., Gagne, G., & Balg, F. (2014). Functional movement screen test: A reliable screening test for young elite ice hockey players. *Physical Therapy in Sport*, *15*(3), 169–175. https://doi.org/10.1016/j.ptsp.2013.10.001

90. Peate, W. F., Bates, G., Lunda, K., Francis, S., & Bellamy, K. (2007). Core strength: a new model for injury prediction and prevention. *Journal of Occupational Medicine and Toxicology (London, England)*, *2*, 3. https://doi.org/10.1186/1745-6673-2-3

91. Premier League Injury Table. (2017). Retrieved from http://www.physioroom.com/news/english_premier_league/epl_injury_table.php

92. Reilly, T., & Gilbourne, D. (2003). Science and football: a review of applied research in the football codes. *Journal of Sports Sciences*, *21*(9), 693.

93. Rey, E., Padron-Cabo, A., Barcala-Furelos, R., & Mecias-Calvo, M. (2016). Effect of High and Low Flexibility Levels on Physical Fitness and Neuromuscular Properties in Professional Soccer Players. *International Journal of Sports Medicine*, *37*(11), 878–883. https://doi.org/10.1055/s-0042-109268

94. Rolls, A., & George, K. (2004). The relationship between hamstring muscle injuries and hamstring muscle length in young elite footballers. *Physical Therapy in Sport*, *5*(4), 179–187. https://doi.org/10.1016/j.ptsp.2004.08.005

95. Schmitt, B., Tim, T., & McHugh, M. (2012). Hamstring injury rehabilitation and prevention of reinjury using lengthened state eccentric training: a new concept. *International Journal of Sports Physical Therapy*, *7*(3), 333–41. http://www.ncbi.nlm.nih.gov/pmc/articles/PMC3362981/

96. S Schneiders, A. G., Davidsson, A., Hörman, E., & Sullivan, S. J. (2011). Functional movement screen normative values in a young, active population. *International Journal of Sports Physical Therapy*, *6*(2), 75–82.

97. Schneiders, A. G., Takemura, M., & Wassinger, C. a. (2009). A prospective epidemiological study of injuries to New Zealand premier club rugby union players. *Physical Therapy in Sport : Official Journal of the Association of Chartered Physiotherapists in Sports Medicine*, *10*(3), 85–90. https://doi.org/10.1016/j.ptsp.2009.05.001

98. Shani, G., Shapiro, A., Oded, G., Dima, K., & Melzer, I. (2017). Validity of the Microsoft Kinect system in assessment of compensatory stepping behavior during standing and treadmill walking. *European Review of Aging and Physical Activity*, *14*(1), 4. https://doi.org/10.1186/s11556-017-0172-8

99. Shultz, R., Anderson, S. C., Matheson, G. O., Marcello, B., & Besier, T. (2013). Test-retest and interrater reliability of the functional movement screen. *Journal of Athletic Training*, *48*(3), 331–6. https://doi.org/10.4085/1062-6050-48.2.11

100. Small, K., McNaughton, L., Greig, M., & Lovell, R. (2010). The effects of multidirectional soccer-specific fatigue on markers of hamstring injury risk. *Journal of Science and Medicine in Sport / Sports Medicine Australia*, *13*(1), 120–5. https://doi.org/10.1016/j.jsams.2008.08.005

101. Smith, C. A., Chimera, N. J., Wright, N. J., & Warren, M. (2013). Interrater and Intrarater Reliability of the Functional Movement Screen. *Journal of Strength and Conditioning Research*, *27*(4), 982–987. https://doi.org/10.1519/JSC.0b013e3182606df2

102. Söderman, K., Alfredson, H., Pietilä, T., & Werner, S. (2001). Risk factors for leg injuries in female soccer players: a prospective investigation during one out-door season. *Knee Surgery, Sports Traumatology, Arthroscopy : Official Journal of the ESSKA*, *9*(5), 313–21. https://doi.org/10.1007/s001670100228

103. Sorani, D. (2016). Influence of sequence of joint movement on hamstring length test results, *International Journal of Physiotherapy, 3*(2), 166–169. https://doi.org/10.15621/ijphy/2016/v3i2/94876

104. Strudwick, A., Reilly, T., & Doran, D. (2002). Anthropometric and fitness profiles of elite players in two football codes. *Journal of Sports Medicine and Physical Fitness,* 42(2):239-42.

105. Tan, H., Wilson, A. M., & Lowe, J. (2008). Measurement of stride parameters using a wearable GPS and inertial measurement unit. *Journal of Biomechanics*, *41*(7), 1398–1406. https://doi.org/10.1016/j.jbiomech.2008.02.021

106. Taylor, W. R., Ehrig, R. M., Duda, G. N., Schell, H., Seebeck, P., & Heller, M. O. (2005). On the influence of soft tissue coverage in the determination of bone kinematics using skin markers. *Journal of Orthopaedic Research*, *23*(4), 726–734. https://doi.org/10.1016/j.orthres.2005.02.006

107. Teyhen, D. S., Shaffer, S. W., Lorenson, C. L., Halfpap, J. P., Donofry, D. F., Walker, M. J., Dugan, J. L., & Childs, J. D. (2012). The Functional Movement Screen: a reliability study. *The Journal of Orthopaedic and Sports Physical Therapy*, *42*(6), 530–40. https://doi.org/10.2519/jospt.2012.3838

108. Thompson, J. A., Tran, A. A., Gatewood, C. T., Shultz, R., Silder, A., Delp, S. L., & Dragoo, J. L. (2016). Biomechanical Effects of an Injury Prevention Program in Preadolescent Female Soccer Athletes. *The American Journal of Sports Medicine*, 45(2), 294-301. https://doi.org/10.1177/0363546516669326

109. Towlson, C., Midgley, A. W., & Lovell, R. (2013). Warm-up strategies of professional soccer players: practitioners' perspectives. *Journal of Sports Sciences*, *31*(13), 1393–401. https://doi.org/10.1080/02640414.2013.792946

110. Van Beijsterveldt, A. M., Van Der Horst, N., Van De Port, I. G., & Backx, F. J. (2013). How effective are exercise-based injury prevention programmes for soccer players?: A systematic review. *Sports Medicine*, *43*(4), 257–265. https://doi.org/10.1007/s40279-013-0026-0

111. Wainner, R. S., Whitman, J. M., Cleland, J. A., & Flynn, T. W. (2007). Regional Interdependence: A Musculoskeletal Examination Model Whose Time Has Come. *Journal of Orthopaedic & Sports Physical Therapy*, *37*(11), 658–660. https://doi.org/10.2519/jospt.2007.0110

112. Waldén, M., Hägglund, M., & Ekstrand, J. (2005). UEFA Champions League study: a prospective study of injuries in professional football during the 2001-2002 season. *British Journal of Sports Medicine*, *39*(8), 542–6. https://doi.org/10.1136/bjsm.2004.014571

113. Waldron, M., Gray, A., Worsfold, P., & Twist, C. (2014). The reliability of Functional Movement Screening (FMS) and in-season changes in physical function and performance among elite rugby league players. *Journal of Strength and Conditioning Research*, 30(4):910–918. https://doi.org/10.1519/JSC.0000000000000270

114. Wang, L., Hu, W., & Tan, T. (2003). Recent developments in human motion analysis. *Pattern Recognition*, *36*(3), 585–601. https://doi.org/10.1016/S0031-3203(02)00100-0

115. Whiteside, D., Deneweth, J. M., Pohorence, M. A., Sandoval, B., Russell, J. R., McLean, S. G., Zernicke, R. F., & Coulet, G. C. (2014). Grading the Functional Movement Screen:a comparison of manual (real-time) and objective methods. *Journal of Strength & Conditioning Research*, *30*(4), 924–933.

116. Witvrouw, E., Danneels, L., Asselman, P., Have, T. D., & Cambier, D. (2003). Muscle Flexibility as a Risk Factor for Developing Muscle Injuries in Male Professional Soccer Players. A prospective study. *American Journal of Sports Medicine,* 31(1):41-6.

117. Woods, C., Hawkins, R. D., Maltby, S., Hulse, M., Thomas, A., & Hodson, A. (2004). The Football Association Medical Research Programme: an audit of injuries in professional football - analysis of hamstring injuries. *British Journal of Sports Medicine*, *38*(1), 36–41. http://articles.sirc.ca/search.cfm?id=S-960973

118. Yeung, S. S., Suen, A. M. Y., & Yeung, E. W. (2009). A prospective cohort study of hamstring injuries in competitive sprinters: preseason muscle imbalance as a possible risk factor. *British Journal of Sports Medicine*, *43*(8), 589–94. https://doi.org/10.1136/bjsm.2008.056283

119. Zhou, H., & Hu, H. (2008). Human motion tracking for rehabilitation-A survey. *Biomedical Signal Processing and Control*, *3*(1), 1–18. https://doi.org/10.1016/j.bspc.2007.09.001

120. Zhou, H., Stone, T., Hu, H., & Harris, N. (2008). Use of multiple wearable inertial sensors in upper limb motion tracking. *Medical Engineering and Physics*, *30*(1), 123–133. https://doi.org/10.1016/j.medengphy.2006.11.010