Are tibial angles measured with inertial sensors useful surrogates for frontal plane projection angles measured using 2-dimensional video analysis during single leg squat tasks? A reliability and agreement study in elite football (soccer) players

Tom Hughes[a,b,*], Richard K. Jones[c], Chelsea Starbuck[c], Jonathan Picot[a], Jamie C. Sergeant[b,d,1], Michael J. Callaghan[a,b,e,1]

[a] Manchester United Football Club, AON Training Complex, Manchester, UK
[b] Arthritis Research UK Centre for Epidemiology, Centre for Musculoskeletal Research, Manchester Academic Health Science Centre, University of Manchester, Manchester, UK
[c] Health Sciences Research Centre, School of Health Sciences, University of Salford, Salford, UK
[d] Centre for Biostatistics, University of Manchester, Manchester Academic Health Science Centre, Manchester, UK
[e] Department of Health Professions, Manchester Metropolitan University, Manchester, UK

ABSTRACT

During single leg squats (SLS), tibial angle (TA) quantification using inertial measurement units (IMU) may offer a practical alternative to frontal plane projection angle (FPPA) measurement using 2-dimensional (2D) video analysis. This study determined: (i) the reliability of IMUs and 2D video analysis for TA measurement, and 2D video analysis for FPPA measurement; (ii) the agreement between IMU TA and both 2D video TA and FPPA measurements during single leg squats in elite footballers. 18 players were tested on consecutive days. Absolute TA (ATA) and relative TA (RTA) were measured with IMUs. ATA and FPPA were measured concurrently using 2D video analysis. Within-session reliability for all measurements varied across days (intraclass correlation coefficient (ICC) range = 0.27–0.83, standard error of measurement (SEM) range = 2.12–6.23°, minimal detectable change (MDC) range = 5.87–17.26°). Between-sessions, ATA reliability was good for both systems (ICCs = 0.70–0.74, SEMs = 1.64–7.53°, MDCs = 4.55–7.01°), while IMU RTA and 2D FPPA reliability ranged from poor to good (ICCs = 0.39–0.72, SEMs = 2.60–5.99°, MDCs = 7.20–16.61°). All limits of agreement exceeded a 5° acceptability threshold. Both systems were reliable for between-session ATA, although agreement was poor. IMU RTA and 2D video FPPA reliability was variable. For SLS assessment, IMU derived TAs are not useful surrogates for 2D video FPPA measures in this population.

1. Introduction

In professional football (soccer), periodic health examination (PHE) is commonly used to obtain performance or rehabilitation benchmarks and to assess potential prognostic factors (predictors) associated with future injuries (Hughes et al., 2018). The single leg squat (SLS) can be utilised in PHE to identify abnormal frontal plane lower extremity kinematics during a functional task (Crossley et al., 2011), which may indicate poor neuromuscular control (Whatman et al., 2011). Typical abnormal kinematics include increased hip medial rotation and adduction, medial tibial rotation and increased foot pronation (Hollman et al., 2009). These movements increase the knee abduction moment and result in medial knee displacement, also known as dynamic valgus (Bell et al., 2008, Petersen et al., 2017) or medial collapse (Powers, 2003). Such kinematics are implicated in non-contact anterior cruciate ligament injury (Koga et al., 2011, Krosshaug et al., 2007, Walden et al., 2015) and in patellofemoral joint dysfunction (Herrington, 2014, Levinger et al., 2007, Nakagawa et al., 2012, Willson and Davis, 2008, Willy et al., 2012).

Dynamic knee valgus kinematics have traditionally been quantified using 3-dimensional (3D) motion analysis systems to estimate the frontal plane knee abduction angle (Gwynne and Curran, 2014,

---

A) Baseline tibial angle for IMU Absolute Tibial Angle (ATA)/Relative Tibial Angle (RTA) measurements (calculated in the ViPerform software).

B) 2D ATA (calculated in Quintic Sports software) and IMU ATA : (calculated from the ViPerform software).

C) IMU RTA (calculated in the ViPerform software).

D) 2D FPPA (calculated in Quintic Sports software).

(caption on next page)

Fig. 1. Description of angles measured with IMU and 2D video systems. Photographs to show: (A) Baseline tibial angle (TA) for IMU absolute tibial angle (ATA)/ relative tibial angle (RTA) measurements, formed between absolute vertical (solid red line) and line from ankle to knee marker (yellow dashed line); (B) 2D ATA formed from the angle between the absolute vertical line (solid red line) and the line from the ankle marker to the knee marker (solid blue line) at the maximal knee flexion angle during the SLS; IMU ATA formed by calculation of sum of baseline TA (angle between dashed yellow line and solid red absolute vertical line), plus the RTA (angle between dashed yellow line and solid blue tibial line) at maximal ankle dorsiflexion during the SLS (Hu et al., 2014); (C) IMU RTA measurement, formed between the baseline TA (yellow dashed line) and TA angle at maximal ankle dorsiflexion (blue solid line) during the SLS; (D) 2D FPPA measurement, calculated by measuring the angle formed between the line from the ASIS marker to knee joint marker (solid green line) and the line from the knee joint marker to the ankle marker (solid blue line), at the frame that corresponded to maximum knee flexion angle during the SLS (Gwynne and Curran, 2014, Herrington et al., 2017, Munro et al., 2012, Willson and Davis, 2008). Please note that the yellow dashed lines in pictures B & C have been adjusted to prevent overlap and aid clarity for the reader. For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.

Table 1
Within-session descriptive statistics.

Measure	System	Day 1				Day 2			
		Preferred leg		Non-preferred leg		Preferred leg		Non-preferred leg	
		Mean (SD)	n	Mean (SD)	n	Mean (SD)	n	Mean (SD)	n
ATA (°)	IMU	9.19 (7.00)	18	13.30 (8.00)	18	8.35 (6.69)	17	11.45 (6.36)	17
	2D Video	9.88 (4.23)	17	11.01 (4.02)	18	9.78 (4.54)	17	10.65 (4.23)	17
RTA (°)	IMU	4.63 (7.30)	18	8.30 (7.68)	18	3.31 (8.60)	17	6.92 (6.68)	17
FPPA (°)	2D Video	1.36 (8.79)	17	4.75 (7.63)	18	2.15 (7.98)	17	3.71 (9.08)	17

Key: SD = standard deviation; n = number of participants; ATA = absolute tibial angle; RTA = relative tibial angle; Note for ATA/RTA measurements + ve values indicate tibial abduction (tibial lateral movement from vertical), −ve values indicate tibial adduction (tibial medial movement from vertical). FPPA = frontal plane projection angle; Note for FPPA measurements + ve values indicate varus alignment, −ve values indicate valgus alignment.

Herrington et al., 2017). These systems are complex and expensive so are limited to use within a laboratory (Hu et al., 2014), which restricts clinical applicability (Willson and Davis, 2008) especially for PHE purposes. A simpler, less expensive alternative has been proposed, where observers use 2-dimensional (2D) analysis software to retrospectively measure femoral and tibial angles from video recordings of SLS performance (Scholtes and Salsich, 2017). These are used to calculate the frontal plane projection angle (FPPA), formed between lines from the anterior superior iliac spine (ASIS) to the knee and from the knee to the ankle at the maximum range of knee flexion (Willson and Davis, 2008, Willson et al., 2006). 2D FPPA has been validated against 3D knee abduction angle measurements in healthy recreational adults (Herrington et al., 2017) and in females with patellofemoral pain (Willson and Davis, 2008). 2D FPPA also has good to excellent within and between-session reliability in recreationally active adults (Gwynne and Curran, 2014, Herrington et al., 2017, Munro et al., 2012). However, population characteristics have considerable influence on reliability (Kottner et al., 2011) and to date, 2D FPPA has not been examined in elite football players (Hughes et al., 2017).

Skin mounted tibial inertial measurement units (IMUs) can also be used to evaluate SLS kinematics (Whelan et al., 2017) and offer practical benefits over 2D systems because real time data analysis may improve time and cost effectiveness. Additionally, IMUs are affordable and portable (Liikavainio et al., 2007) so could be more useful than 3D systems in clinical environments (Charry et al., 2013). IMU systems use data from integrated accelerometers, magnetometers and gyroscopes to estimate tibial angles (TA) in the frontal plane during SLS tasks, which appear to correspond with FPPA values (Hu et al., 2014).

It is important to evaluate the reliability and agreement of alternative systems compared to established clinically relevant methods before incorporating their use in practice (Luiz et al., 2003). While the reliability of IMU systems for TA measurement has not been examined in elite footballers, neither has the agreement between IMU derived TA measurements and 2D video FPPA measurements. Hence it is unclear if an IMU system could be used as a surrogate method of 2D FPPA measurement. The aims of this study were to determine during a SLS in elite footballers: (i) the reliability of an IMU and a 2D video system for TA measurement and a 2D video system for FPPA measurement; (ii) the agreement between IMU TA measurements and 2D video FPPA and TA

measurements.

2. Materials and methods

This study was conducted and is reported in accordance with the Guidelines for Reporting Reliability and Agreement Studies (Kottner et al., 2011).

2.1. Participants

A convenience sample of 18 participants was selected from a cohort of elite male football players under contract at an English Premier League Football Club. Informed consent was not required because all data were captured from mandatory PHE processes completed through the participant's employment. The anonymity and rights of all participants were protected. The football club granted permission to use these data. The use of these data for the current purpose was approved by the Research Ethics Service at the University of Manchester.

2.2. Eligibility criteria

Participants were included if they: (i) were > 16 years and < 40 years old; (ii) trained fully without injury and available for match selection within two weeks of testing. Participants were excluded if they: (i) were a goalkeeper; (ii) had undergone previous major lower extremity joint surgery; (iii) had a true leg length discrepancy of > 1 cm (cm); iv) suffered a systemic illness within the week before testing.

2.3. Preparation

Baseline measurements were recorded of: (i) standing height (cm) and body mass (kilograms) using a scale and height measure (SECA 220, SECA, Hamburg, Germany); (ii) true leg length for each limb using a cloth tape measure (Magee, 2008); (iii) the participants' preference for kicking and non-kicking leg. Participants were instructed to wear the same footwear and use orthotics if previously prescribed.

The IMU used was the ViPerform system (Dorsavi, Melbourne, Australia) which consisted of two 3D IMUs sampling at 100, 20 and 20 Hz on the x, y and z axes respectively (Charry et al., 2013). IMUs

Table 2
Within-session reliability ICCs $_{3,1}$, 95% CI, SEM and MDC values for both systems and days, per limb preference.

| Parameter | System | Day 1 | | | | | | | | Day 2 | | | | | | | |
| | | Preferred leg | | | | Non preferred leg | | | | Preferred leg | | | | Non preferred leg | | | |
		ICC	95% CI	SEM (°)	MDC (°)	ICC	95% CI	SEM (°)	MDC (°)	ICC	95% CI	SEM (°)	MDC (°)	ICC	95% CI	SEM (°)	MDC (°)
ATA	IMU	0.30	0.00–0.62	5.87	16.26	0.53	0.25–0.76	5.50	15.25	0.59	0.31–0.81	4.29	11.89	0.65	0.39–0.84	3.77	10.46
	2D Video	0.74	0.52–0.89	2.15	5.95	0.54	0.26–0.77	2.73	7.57	0.50	0.21–0.76	3.21	8.89	0.75	0.53–0.89	2.12	5.87
RTA	IMU	0.27	−0.01 to 0.59	6.23	17.26	0.48	0.19–0.73	5.54	15.37	0.75	0.54–0.89	4.29	11.88	0.62	0.36–0.83	4.10	11.36
FPPA	2D Video	0.83	0.67–0.93	3.61	10.01	0.53	0.25–0.77	5.24	14.52	0.63	0.36–0.83	4.87	13.50	0.81	0.64–0.92	3.93	10.90

Key: ICC = Intraclass correlation coefficient; CI = confidence interval; SEM = standard error of measurement, MDC = minimal detectable change; ATA = absolute tibial angle; RTA = relative tibial angle; Note for ATA/ RTA measurements +ve values indicate tibial abduction (tibial lateral movement from vertical), −ve values indicate tibial adduction (tibial medial movement from vertical; +ve measurements indicate varus alignment, −ve measurements indicate valgus alignment.

were applied according to the manufacturer's instructions using a proprietary leg template to identify the correct site, based on each participant's height. Disposable application pads (Dorsavi, Melbourne, Australia) were affixed to the medial tibia and the IMU units were clipped into position. For the 2D system, one video camera (Sanyo Xacti, Sanyo, Osaka, Japan) was located 3 m from the participant on a 60 cm high tripod. Zinc oxide tape markers were placed on FPPA landmarks bilaterally at the midpoints of the ankle malleoli, femoral condyles and ASIS (Willson and Davis, 2008, Willson et al., 2006). Equipment application was completed by a physiotherapist (JP), experienced in IMU and video analysis.

2.4. Data capture

IMU data and 2D video capture was completed concurrently by the same physiotherapist (JP). IMU data were analysed in real time and saved to a computer using the manufacturer's software (ViPerform 5.10, Dorsavi, Melbourne, Australia). Video footage was digitised and analysed retrospectively using Quintic Sports software (Quintic Sports, Quintic Consultancy Ltd, Sutton Coldfield, UK) by a post-doctoral biomechanist (CS) experienced in 2D video analysis.

2.5. Measurement parameters

All parameters are shown in Fig. 1. Because absolute tibial angle (ATA) was a component of both IMU and 2D FPPA analysis, this was directly compared. Relative tibial angle (RTA) was provided by the IMU system only. For all TA measurements, positive values indicated tibial abduction (proximal tibial lateral displacement from vertical) and negative values indicated tibial adduction (proximal tibial medial displacement from vertical). 2D FPPA was calculated using the video system. Positive values represented varus knee alignment, whereas negative values represented valgus alignment. FPPA could not be calculated directly using the IMU system, therefore IMU ATA and RTA measurements were compared as surrogate FPPA measures.

2.6. Procedure

Testing sessions were performed before the players commenced training, at least 2 days following game participation to reduce potential fatigue effects. To warm up, participants cycled for 5 minutes on an exercise bicycle without resistance and completed three practice SLS attempts on each leg. The standardised start position is presented in Fig. 1(A). From this position, participants were instructed to perform the SLS to at least 45° and no greater than 60° knee flexion (Zeller et al., 2003) over 5 s using a timer, which was verbalised by the examiner (JP); second 1 initiated the SLS, second 3 indicated the position of maximal knee flexion and second 5 indicated the trial end, where ground contact with both legs was permitted (Herrington, 2014). This standardisation reduced any velocity effects on kinematics. Knee flexion range was controlled for through examiner observation and feedback. Data were captured for 5 trials per leg, with the left side evaluated before the right side. Data were analysed from trials 3, 4 and 5 only. If a trial was discounted due to inappropriate technique or a technical problem, either trial 1 or 2 was included as an alternative. The same procedure was repeated 24 h later, to minimise any effects from football training.

2.7. Data analysis

Within-session analyses were conducted for both testing days, stratified by kicking limb preference. For between-session analyses, mean values for the 3 trials were calculated for each testing day and participant, per kicking limb preference. Within and between-session ICCs$_{3,1}$ with corresponding 95% confidence intervals (CIs) were calculated and interpreted as: poor < 0.40; fair = 0.40–0.70;

Table 3
Between-session Means, SD, ICCs$_{3,1}$, SEM and MDC values for both systems, per limb preference.

Parameter	System	Preferred leg						Non preferred leg					
		Mean (SD) (°)	n	ICC	95% CI	SEM (°)	MDC (°)	Mean (SD) (°)	n	ICC	95% CI	SEM (°)	MDC (°)
ATA	IMU	6.70 (4.39)	17	0.70	0.35–0.88	2.39	6.63	11.85 (4.97)	17	0.74	0.42–0.90	2.53	7.01
	2D Video	10.19 (3.19)	16	0.74	0.39–0.90	1.64	4.55	10.94 (3.30)	17	0.73	0.39–0.89	1.73	4.78
RTA	IMU	2.82 (6.07)	17	0.55	0.10–0.81	4.09	11.35	7.32 (5.41)	17	0.77	0.47–0.91	2.60	7.20
FPPA	2D Video	1.55 (7.69)	16	0.39	−0.11–0.74	5.99	16.61	4.90 (7.33)	17	0.74	0.42–0.90	3.73	10.33

Key: SD = standard deviation; n = number of participants; ICC = Intraclass correlation coefficient; CI = confidence interval, SEM = standard error of measurement, MDC = minimal detectable change; ATA = absolute tibial angle; RTA = relative tibial angle; Note for ATA/RTA measurements +ve values indicate tibial abduction (tibial lateral movement from vertical), −ve values indicate tibial adduction (tibial medial movement from vertical; FPPA = frontal plane projection angle, +ve measurements indicate varus alignment, −ve measurements indicate valgus alignment.

good = 0.70–0.90; excellent ≥ 0.90 (Coppieters et al., 2002). For all parameters, standard error of measurement (SEM) and minimal detectable changes (MDC) were calculated as described by Weir (2005). MDCs were considered acceptable if values were < 10°.

Because ATA could be compared from both the IMU and the 2D video systems, Bland–Altman plots with 95% limits of agreement (LOA) were produced (Bland and Altman, 1986) according to limb preference for each testing day. This process was repeated to establish the agreement between IMU ATA or RTA and 2D FPPA measures. Agreement was considered acceptable if the 95% LOA fell within a 5° range, as described previously (Gwynne and Curran, 2014). Statistical analyses were completed using STATA 14 (StataCorp LLC, Texas, USA).

3. Results

3.1. Study participation and missing data

The characteristics of the sample were (expressed as mean ± standard deviations): age 17.45 ± 1.14 years; height 180.32 ± 6.85 cm; weight 74.32 ± 5.62 kg; preferred kicking leg length 99.19 ± 4.93 cm; non-preferred kicking leg length 99.18 ± 4.87 cm. Orthotics were used in the training shoes of 6 participants for both testing sessions, as they had been previously prescribed by a podiatrist. A technical fault affected 2D ATA and FPPA data from the preferred leg of one participant on Day 1, which was excluded from the relevant within-session analysis. One participant was injured before Day 2 testing and was excluded from the relevant within-session analysis. In both cases, these data were excluded from the between-session analysis.

3.2. Within-session reliability

Within-session descriptive statistics are presented in Table 1 while ICCs and 95% CIs, SEM and MDC values are summarised in Table 2. For ATA, IMU ICCs ranged from poor (0.30) to fair (0.65) while the 2D video ICCs ranged from fair (0.50) to good (0.75) across legs and days. SEMs and MDCs were lowest overall for 2D video ATA (range = 2.12–3.21° and 5.87–8.89° respectively), but larger for IMU ATA (range = 3.77–5.87° and 10.46–16.26° respectively). For RTA, IMU ICCs ranged from poor (0.27) to good (0.75), SEMs ranged from 4.10° to 6.23° and MDCs ranged from 11.36° to 17.26°. For 2D FPPA, ICCs ranged from fair to good (0.53–0.83), SEMs ranged from 3.61 to 5.24° and MDCs ranged from 10.01° to 14.52°. Other than 2D ATA and 2D FPPA on the preferred kicking leg, all other reliability measurements improved on Day 2.

3.3. Between-session reliability

Between-session reliability statistics are presented in Table 3. For ATA, ICCs were good bilaterally for both systems (range 0.70–0.74). SEMs and MDCs were larger for the IMU (2.39–2.53°, 6.63–7.01°

respectively) compared to the 2D system (1.64–1.73°, 4.55–4.78° respectively) but were similar across limbs. For RTA, IMU ICCs were fair (0.55) on the preferred kicking leg, but good on the non-preferred leg (0.77). SEMs and MDCs were larger on the preferred leg (4.09° and 11.35°) compared to the non-preferred leg (2.60° and 7.20°). 2D FPPA ICCs were poor (0.39) on the preferred leg, compared to good for the non-preferred leg (0.74). Similarly, SEM and MDCs were greater on the preferred leg (5.99° and 16.61° respectively) compared to the non-preferred leg (3.73° and 10.33°).

3.4. Between-system agreement

Bland-Altman plots for each day are presented in Figs. 2 and 3. The observed mean agreement for: (i) ATA between systems ranged from −1.5° to 2.29°; (ii) IMU RTA and 2D FPPA ranged from 2.08° to 3.54°; (iii) IMU ATA and 2D FPPA ranged from 7.10° to 8.54°. For all measures, the 95% LOA extended beyond the 5° threshold set *a priori,* which indicated unacceptable agreement.

4. Discussion

This study has determined during the SLS test in elite footballers: (i) the reliability of an IMU and 2D video system for TA measurement and a 2D video system for FPPA measurement; (ii) the agreement between IMU TA measurements and 2D TA and FPPA measurements. In taking these assessments we wanted to answer the question, are tibial angles measured with IMUs useful surrogate measures of frontal plane projection angles during SLS tasks? Our results show that this is not the case.

4.1. Tibial angles

Despite generally high levels of between participant variability for ATA and RTA parameters, irrespective of the system used, we observed that on average elite footballers demonstrated tibial *abduction* angles during the SLS task which were associated with knee *varus* alignment. Interestingly, both within and between-session mean ATA and RTA values were consistently larger in the non-preferred kicking leg.

This is the only study known to investigate TA measurement reliability and cannot be compared with other work. Our within-session analyses have shown that both systems suffered from variable reliability for ATA measurement (IMU ICC range = 0.30–0.65, 2D video ICC range = 0.50–0.75) and 2D video ATA was the only parameter to have clinically acceptable MDC values. Therefore, using either system for within-session measurement of ATA is unlikely to be clinically useful in this population and any results should be interpreted with caution. Despite this, either system could be useful in clinical practice to monitor between-session changes of ATA, because both systems have comparably good reliability (IMU ICC range = 0.70–0.74, 2D video ICC range = 0.73–0.74) and averaging the scores for each session reduces the data variability and measurement error. As the IMU and 2D video
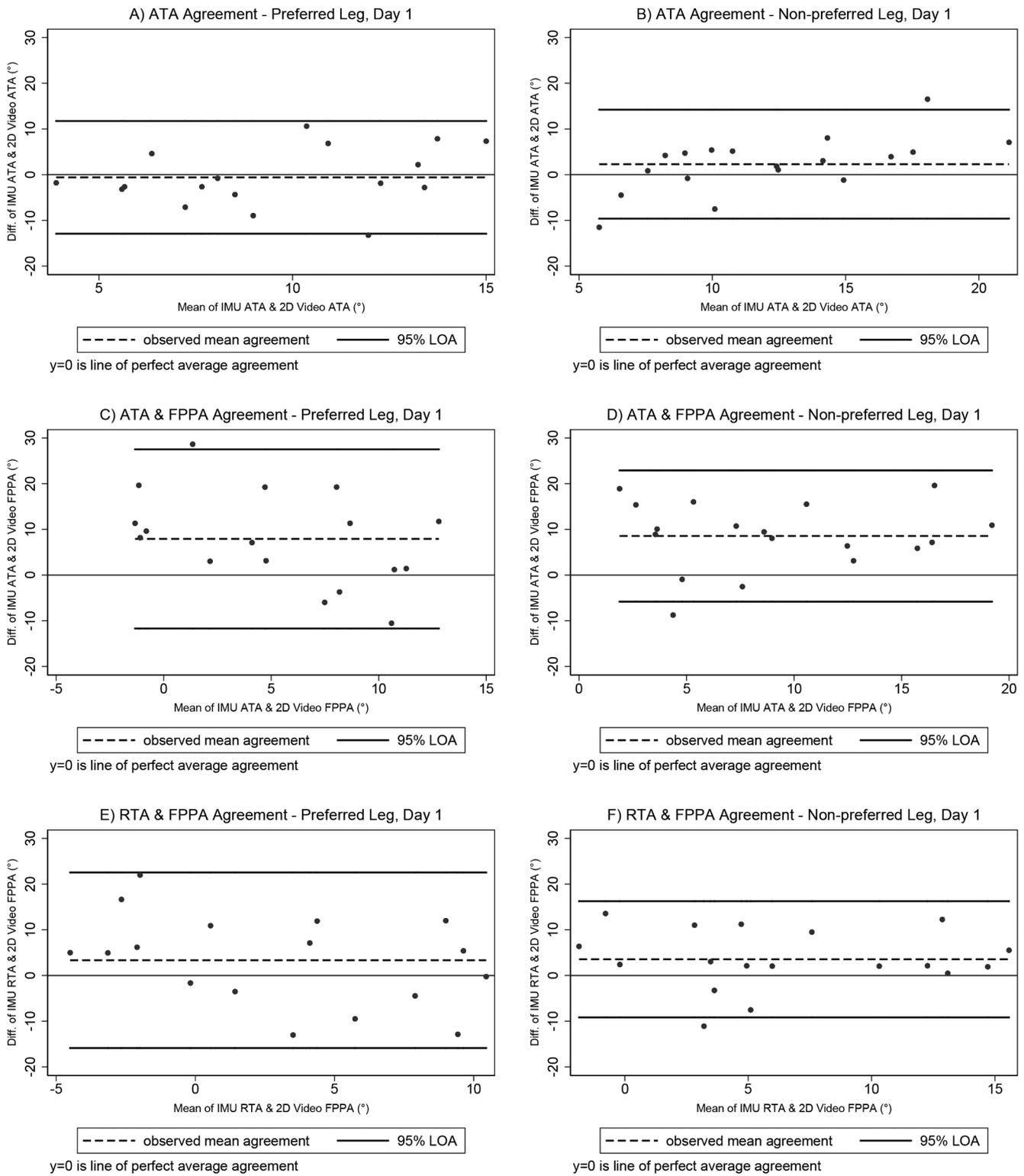
Fig. 2. Agreement between measurement parameters and systems on Day 1. Key: ATA = absolute tibial angle; RTA = relative tibial angle; FPPA = frontal plane projection angle; Diff = difference, IMU = inertial measurement unit; 2D = 2 dimensional; LOA = limits of agreement; ° = degrees.
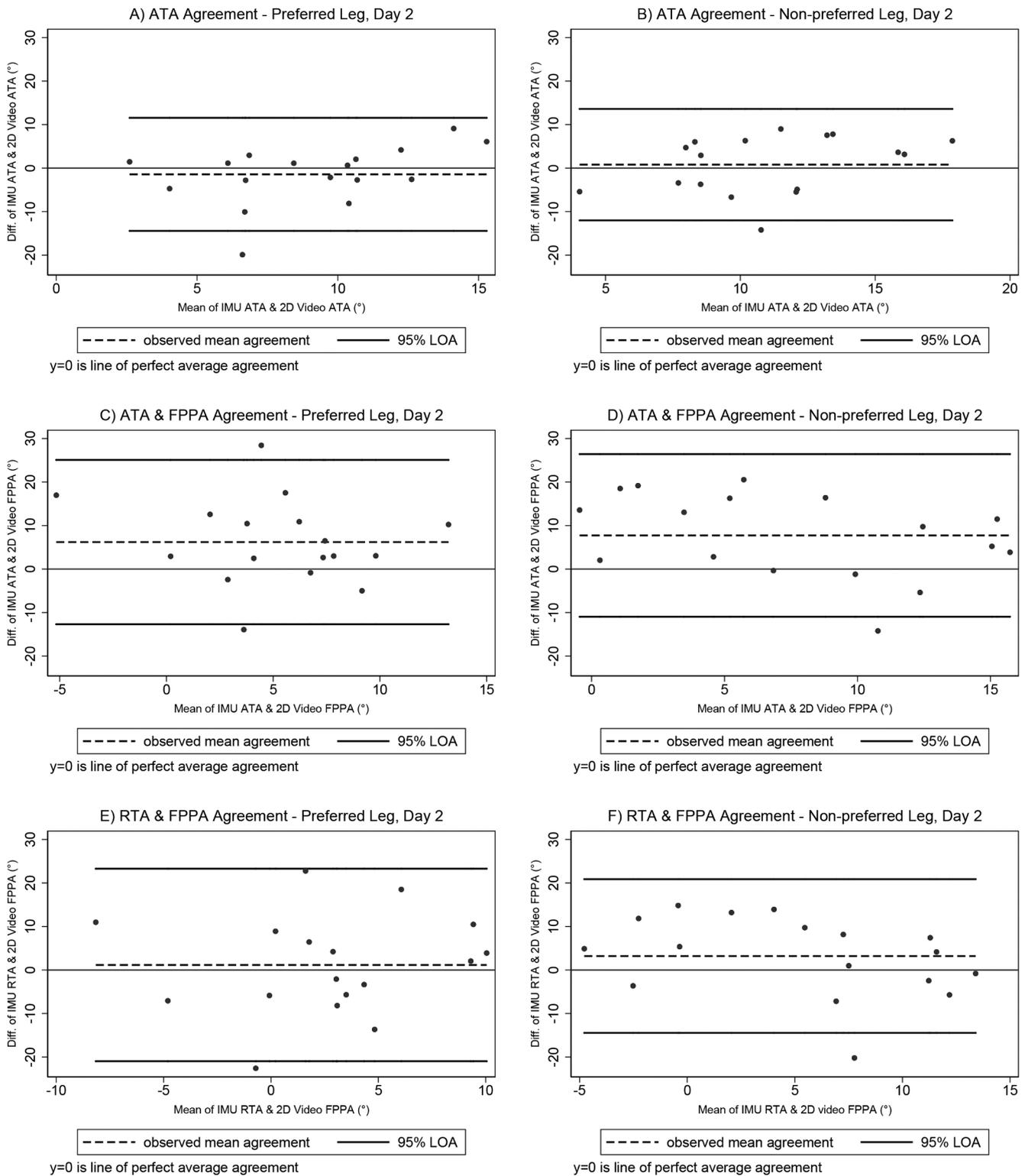
Fig. 3. Agreement between measurement parameters and systems on Day 2. Key: ATA = absolute tibial angle; RTA = relative tibial angle; FPPA = frontal plane projection angle; Diff = difference, IMU = inertial measurement unit; 2D = 2 dimensional; LOA = limits of agreement; ˚ = degrees.

systems did not have acceptable agreement for ATA measures, ATA should be considered as a system specific, arbitrary measure of tibial motion. However, the relevance of ATA as a performance indicator, rehabilitation marker or potential prognostic factor for injury is unknown at present, so further research is required to establish the clinical usefulness of this measurement parameter.

RTA measurement reliability of the IMU system was mostly inferior to the related ATA measurements and could not be recommended in practice.

4.2. 2D FPPA

In this study, 2D FPPA measurements were also observed in a *varus* direction. The ranges of within and between-session mean values were 1.36–2.15° for the preferred leg, and 3.71–4.90° for the non-preferred leg. This apparent effect of limb preference on kinematics has not been observed previously. Our results differ from the *valgus* direction (mean FPPA = 8.64° SD = 9.06°) observed by Munro et al., 2012) and Gwynne and Curran (2014) (mean FPPA = 7.80°, SD = 7.33), although are more comparable to the *varus* direction observed by Herrington et al. (2017) (mean FPPA range from 9.1 (SD = 10.6) to 11.7 (SD = 9.8) by two separate raters). These kinematic differences may be due to population characteristics, as none these studies used elite sports people. Between-participant variability in our study was comparable for both limbs and confirms previous work (Gwynne and Curran, 2014, Herrington et al., 2017, Munro et al., 2012). This suggests that large kinematic variability appears to be trait of 2D FPPA measurement, regardless of population.

In recreationally active adults, 2D FPPA has been shown to have good to excellent within-session reliability (ICCs 0.72–0.86) and between-session reliability (ICCs 0.74–0.87) (Gwynne and Curran, 2014, Herrington et al., 2017, Munro et al., 2012). We found that in elite footballers, within-session reliability for 2D FPPA was uncertain because of ICC variability between testing days, which ranged from fair to good (0.53–0.83). This has also not been observed previously. We also found generally greater within-session SEM values in our population (range = 3.61–5.24°) compared to values reported previously (range = 1.72°–2.10°) (Gwynne and Curran, 2014, Herrington et al., 2017) which may account for the clinically unacceptable MDC values demonstrated (range = 10.01–14.52°).

During between-session analysis, we found that 2D FPPA reliability was dependent on limb kicking preference. On the non-preferred leg, between-session ICCs were good (0.74) and replicated ICC results previously reported by Gwynne and Curran (2014) although were less than those cited by Munro et al. (2012) (ICC = 0.88) and (Herrington et al., 2017) (ICC = 0.87). However, we found that for the preferred kicking leg, between-session reliability was poor (0.39). This may be due to differences in SEM values observed between limbs, where the non-preferred leg was 3.73° and comparable to previous work (range = 1.37°–3.82°) (Gwynne and Curran, 2014, Herrington et al., 2017, Munro et al., 2012), whereas the SEM for the preferred leg (5.99°) was greater than previously reported. Subsequently for either limb, the minimal differences required in between-session 2D FPPA performances are too large (10.33–16.01°) to be helpful in detecting real changes in elite footballers. In comparison, 2D FPPA MDCs have been found to be 7.63° and 8.93° in recreationally active men and women (Munro et al., 2012) but these differences may be due to characteristics of the elite football players investigated in our study.

The kicking action in football places differing demands on the support limb and kicking limb musculature (Brophy et al., 2007), so kicking limb preference may be associated with specific musculoskeletal adaptations from training exposure. To account for this, our analyses were stratified according to limb kicking preference, whereas previously 2D FPPA data from both legs were either pooled (Herrington et al., 2017, Munro et al., 2012) or it was unclear which limbs were evaluated (Gwynne and Curran, 2014). This may partially explain the differences in kinematics, reliability and error measurements observed in our study compared to those previously reported. We also controlled for biomechanical foot differences through standardised use of corrective orthoses (if prescribed), which may have also have influenced 2D FPPA measurement and the degree of measurement error. It was unclear if this was controlled for in previous studies (Gwynne and Curran, 2014, Herrington et al., 2017, Munro et al., 2012).

Considering our results, we would argue that in this population, using video analysis to measure 2D FPPA is inadequate as a cross sectional assessment or as a rehabilitation/PHE test where performance could be monitored longitudinally. We have also shown that absolute and relative tibial angles measured using IMUs did not agree sufficiently with 2D FPPA, so could not be considered as valid surrogate measures. This is unsurprising because femoral and pelvic measurements are required for FPPA quantification (Willson and Davis, 2008) and these angles are not recorded with tibial mounted IMUs.

4.3. Limitations and future research

The general improvement of all TA reliability statistics during the second session is suggestive of learning effects, although this was surprising as practice trials were permitted. Order bias was partially controlled for through analysis stratification by limb preference, although limbs were tested on the left and right side consecutively and not randomised, which could have influenced the results. The generalisability of our findings is limited to elite football players and is specific to analysis of SLS tasks only.

Further research could involve replication studies which randomise the limb testing order, using participants from different sports or non-elite football populations, and evaluation of frontal plane control during other dynamic functional tasks.

5. Conclusion

This study demonstrates the reliability limitations of an IMU and 2D video analysis system for SLS kinematic assessment in elite football players, especially during within-session analyses. For between-session analysis of ATA, both systems were sufficiently reliable to monitor SLS performance although agreement was poor and the clinical relevance of ATA is unknown at present. The varied between-session reliability of IMUs for RTA measurement and 2D video for FPPA measurement means that in elite players, the value of using these parameters to quantify SLS kinematics in clinical practice is questionable. Furthermore, in this population TA measured with IMUs cannot be considered as surrogate FPPA measures of SLS tasks due to inadequate agreement between systems. It should be remembered that this study has exclusively investigated SLS performance. To firmly establish the usefulness of IMU and 2D video systems as clinical assessment tools, evaluation of alternative functional tasks is recommended.

Conflict of interest

None declared.

Ethical approval

Informed consent was not required; all data were captured from mandatory PHE processes completed through the participant's employment. The use of these data for the current purpose was approved by the Research Ethics Service at the University of Manchester.

Data sharing statement

The participating football club granted permission to use the collected data for research purposes. An anonymised summary of the dataset generated and analysed during the current study may be available from the corresponding author on reasonable request.

Acknowledgements

The authors would like to thank all staff within the medical department at Manchester United for their support and Dorsavi for their technical advice.

References

Bell, D.R., Padua, D.A., Clark, M.A., 2008. Muscle strength and flexibility characteristics of people displaying excessive medial knee displacement. Arch. Phys. Med. Rehabil. 89, 1323–1328.

Bland, J.M., Altman, D.G., 1986. Statistical methods for assessing agreement between two methods of clinical measurement. Lancet. 346, 307–310.

Brophy, R.H., Backus, S.I., Pansy, B.S., Lyman, S., Williams, R.J., 2007. Lower extremity muscle activation and alignment during the soccer instep and side-foot kicks. J. Orthop. Sports Phys. Ther. 37, 260–268.

Charry, E., Hu, W., Umer, M., Ronchi, A., Taylor, S., 2013. Study on estimation of peak ground reaction forces using tibial accelerations in running. Intelligent Sensors, Sensor Networks and Information Processing, IEEE Eighth International Conference. Melbourne, Australia.

Coppieters, M., Stappaerts, K., Janssens, K., Jull, G., 2002. Reliability of detecting 'onset of pain' and 'submaximal pain' during neural provocation testing of the upper quadrant. Physiother. Res. Int. 7, 146–156.

Crossley, K.M., Zhang, W.-J., Schache, A.G., Bryant, A., Cowan, S.M., 2011. Performance on the single-leg squat task indicates hip abductor muscle function. Am. J. Sports Med. 39, 866–873.

Gwynne, C.R., Curran, S.A., 2014. Quantifying frontal plane knee motion during single limb squats: reliability and validity of 2-dimensional measures. Int. J. Sports Phys. Ther. 9, 898–906.

Herrington, L., 2014. Knee valgus angle during single leg squat and landing in patellofemoral pain patients and controls. Knee 21, 514–517.

Herrington, L., Alenezi, F., Alzhrani, M., Alrayani, H., Jones, R., 2017. The reliability and criterion validity of 2D video assessment of single leg squat and hop landing. J. Electromyogr. Kinesiol. 34, 80–85.

Hollman, J.H., Ginos, B.E., Kozuchowski, J., Vaughn, A.S., Krause, D., Youdas, J., 2009. Relationships between knee valgus, hip-muscle strength, and hip-muscle recruitment during a single-limb step-down. J. Sport Rehabil. 18, 104–117.

Hu, W., Charry, E., Umer, M., Ronchi, A., Taylor, S., 2014. An inertial sensor system for measurements of tibia angle with applications to knee valgus/varus detection. Intelligent Sensors, Sensor Networks and Information Processing, Ninth IEEE International Conference. Singapore.

Hughes, T., Sergeant, J.C., Parkes, M., Callaghan, M.J., 2017. Prognostic factors for specific lower extremity and spinal musculoskeletal injuries identified through medical screening and training load monitoring in professional football (soccer): a systematic review. BMJ Open Sport Exerc. Med. 3, 1–18.

Hughes, T., Sergeant, J.C., van der Windt, D.A., Riley, R., Callaghan, M.J., 2018. Periodic health examination and injury prediction in professional football (Soccer): theoretically, the prognosis is good. Sports Med.

Koga, H., Bahr, R., Myklebust, G., Engebretsen, L., Grund, T., Krosshaug, T., 2011. Estimating anterior tibial translation from model-based image-matching of a noncontact anterior cruciate ligament injury in professional football: a case report. Clin. J. Sport Med. 21, 271–274.

Kottner, J., Audige, L., Brorson, S., Donner, A., Gajewski, B.J., Hrobjartsson, A., et al., 2011. Guidelines for reporting reliability and agreement studies (GRRAS) were proposed. Int. J. Nurs. Stud. 48, 661–671.

Krosshaug, T., Nakamae, A., Boden, B.P., Engebretsen, A.H., Smith, G., Slauterbeck, J.R., et al., 2007. Mechanisms of anterior cruciate ligament injury in basketball: video analysis of 39 cases. Am. J. Sports Med. 35, 359–367.

Levinger, P., Gilleard, W., Coleman, C., 2007. Femoral medial deviation angle during a one-leg squat test in individuals with patellofemoral pain syndrome. Phys. Ther. Sport. 8, 163–168.

Liikavainio, T., Bragge, T., Hakkarainen, M., Jurvelin, J.S., Karjalainen, P.A., Arokoski, J.P., 2007. Reproducibility of loading measurements with skin-mounted accelerometers during walking. Arch. Phys. Med. Rehabil. 88, 907–915.

Luiz, R., Costa, A.J.L., Kale, P.L., Werneck, G.L., 2003. Assessment of agreement of a quantitative variable: a new graphical approach. J. Clin. Epidemiol. 56, 963–967.

Magee, D.J., 2008. Orthopedic Physical Assessment, fifth ed. Saunders Elsevier, Missouri.

Munro, A., Herrington, L., Carolan, M., 2012. Reliability of 2-dimensional video assessment of frontal plane dynamic knee valgus during common athletic screening tasks. J. Sport Rehabil. 21, 7–11.

Nakagawa, T.H., Moriya, E.T.U., Maciel, C.D., Serrao, F.V., 2012. Trunk, pelvis, hip, and knee kinematics, hip strength, and gluteal muscle activation during a single-leg squat in males and females with and without patellofemoral pain syndrome. J. Orthop. Sports Phys. Ther. 42, 491–501.

Petersen, W., Rembitzki, I., Liebau, C., 2017. Patellofemoral pain in athletes. Open Access J. Sports Med. 8, 143–154.

Powers, C.M., 2003. The influence of altered lower extremity kinematics on patellofemoral joint dysfunction: a theoretical perspective. J. Orthop. Sports Phys. Ther. 33, 639–646.

Scholtes, S.A., Salsich, G.B., 2017. A dynamic valgus index that combines hip and knee angles: assessment of utility in females with Patellofemoral pain. Int. J. Sports Phys. Ther. 12, 333–340.

Walden, M., Krosshaug, T., Bjorneboe, J., Andersen, T.E., Faul, O., Hagglund, M., 2015. Three distinct mechanisms predominate in non-contact anterior cruciate ligament injuries in male professional football players: a systematic video analysis of 39 cases. Br. J. Sports Med. 49, 1452–1460.

Weir, J.P., 2005. Quantifying test-retest reliability using the intraclass correlation coefficient and the SEM. J. Strength Condition. Res. 19, 231–240.

Whatman, C., Hing, W., Hume, P., 2011. Kinematics during lower extremity functional screening tests–are they reliable and related to jogging? Phys. Ther. Sport 12, 22–29.

Whelan, D.F., O'Reilly, M.A., Ward, T.E., Delahunt, E., Caulfield, B., 2017. Technology in rehabilitation: evaluating the single leg squat exercise with wearable inertial measurement units. Methods Inform. Med. 56, 88–94.

Willson, J.D., Davis, I., 2008. Utility of the frontal plane projection angle in females with Patellofemoral pain. J. Orthop. Sports Phys. Ther. 38, 606–615.

Willson, J.D., Ireland, M.L., Davis, I., 2006. Core strength and lower extremity alignment during single leg squats. Med. Sci. Sports Exerc. 38, 945–952.

Willy, R.W., Manal, K.T., Witvrouw, E.E., Davis, I.S., 2012. Are mechanics different between male and female runners with patellofemoral pain? Med. Sci. Sports Exerc. 44, 2165–2171.

Zeller, B.L., McCrory, J.L., Kibler, W.B., Uhl, T.L., 2003. Differences in kinematics and electromyographic activity between men and women during the single-legged squat. Am. J. Sports Med. 31, 449–468.

Tom Hughes is a Senior Physiotherapist and Clinical Lead for the Professional Development Squad at Manchester United. He is currently completing a PhD at the University of Manchester, investigating the use of periodic health examination and injury prediction/prognosis in elite football and has published other articles in peer-reviewed sports medicine journals in this area.

Richard Jones is a Professor of Clinical Biomechanics at the University of Salford. Prof Jones leads the Knee Biomechanics research programme at the University researching all aspects related to healthy and pathological function and rehabilitation.

Chelsea Starbuck is a Post-Doctoral Research Fellow at the University of Salford. Chelsea's research includes lower body biomechanics during sport specific movements, shoe-surface interactions and biomechanical factors associated with ankle and knee injuries and the development of knee osteoarthritis. Chelsea completed her PhD at the University of Exeter in 2015. Her PhD was a collaborative project with the International Tennis Federation working with existing projects to understanding players' perceptions and biomechanical response to tennis surfaces to provide recommendations and validations for the development of mechanical tests to better characterise tennis surfaces. Chelsea previously lectured at the University of St Mark and St John. Further information on previous research can be found at https://www.seek.salford.ac.uk/profiles/CStarbuck.jsp

Jonathan Picot is a First Team Physiotherapist at Manchester United Football Club. He holds a BSc in Sports Science from Cheltenham and Gloucester College of Higher Education, a BSc in Physiotherapy from the University of Bradford, and an MSc in Advancing Physiotherapy from the University of Salford. He has interests in research relating to athletic performance and advancements in Physiotherapeutic modalities.

Michael Callaghan is Professor of Clinical Physiotherapy at Manchester Metropolitan University and Head of Physical Therapies at Manchester United FC.

Jamie Sergeant is a Lecturer in Biostatistics at the University of Manchester's Centre for Biostatistics. His research is focused on the epidemiology of musculoskeletal diseases and he collaborates closely with colleagues in the Arthritis Research UK Centre for Epidemiology at Manchester. Jamie has a keen interest in the teaching and learning of statistics and how statistics can be communicated effectively to all audiences.