



**Henderson, F. and Alho, R. and Riches, P. and Picard, F. (2017)  
Assessment of knee alignment with varus and valgus force through the  
range of flexion with non-invasive navigation. Journal of Medical  
Engineering and Technology. pp. 1-16. ISSN 0309-1902 ,  
<http://dx.doi.org/10.1080/03091902.2017.1333164>**

This version is available at <https://strathprints.strath.ac.uk/60918/>

**Strathprints** is designed to allow users to access the research output of the University of Strathclyde. Unless otherwise explicitly stated on the manuscript, Copyright © and Moral Rights for the papers on this site are retained by the individual authors and/or other copyright owners. Please check the manuscript for details of any other licences that may have been applied. You may not engage in further distribution of the material for any profitmaking activities or any commercial gain. You may freely distribute both the url (<https://strathprints.strath.ac.uk/>) and the content of this paper for research or private study, educational, or not-for-profit purposes without prior permission or charge.

Any correspondence concerning this service should be sent to the Strathprints administrator: [strathprints@strath.ac.uk](mailto:strathprints@strath.ac.uk)

## **Assessment of knee alignment with varus and valgus force through the range of flexion with non-invasive navigation**

### **Abstract**

In image-free total knee arthroplasty (TKA) navigation, infrared markers are attached to bony landmarks to provide kinematic data intra-operatively, with the aim of improving the precision of implant placement. In non-invasive navigation, infrared markers are attached to the skin surface, with recent evidence suggesting that this can give repeatable measurements of lower limb mechanical alignment. The aim of our study was to evaluate the use of a non-invasive navigation system in the assessment of mechanical alignment with applied coronal force through the range of flexion.

A previously validated non-invasive system (Physiopilot™) was tested on 23 volunteers with healthy knees. Two users performed two registrations of the software workflow on each participant's right and left knees. A force was manually applied to the end-point of varus and valgus knee laxity and the measured change in mechanical alignment was recorded. Force was applied with the knee positioned in increments of flexion from 0°-90°.

In keeping with previous studies, satisfactory values of Coefficient of Repeatability (CR) of 1.55 and 1.33 were found for intra-observer repeatability in measurement of supine Mechanical Femoro-tibial Angle (MFTA) in extension, with a good inter-observer correlation of Intraclass Correlation Coefficient (ICC) 0.72. However, when flexion was introduced, intra-observer and inter-observer reliability fell outwith acceptable limits. The trial therefore did not support use of the Physiopilot™ system as a measure of MFTA when flexion is introduced. It was felt that learning-curve, soft tissue artefacts and lack of force standardisation equipment may have accounted for significant levels of

error, with further studies required to address these issues.

## **Keywords**

Total knee arthroplasty, navigation, CAOS, MFTA, non-invasive

## **Introduction**

Osteoarthritis (OA) is a highly prevalent disease, with approximately 13% of patients over the age of 55 experiencing symptomatic knee osteoarthritis [1]. The knee is the most commonly affected large joint in the body [1], and OA remains the major indication for total knee arthroplasty (TKA) [2].

A number of studies have pointed to the significance of coronal mechanical alignment of the knee to the outcome of TKA. The mechanical femoro-tibial angle (MFTA), as assessed in this study (Figure 1), is the angle in the coronal plane formed between the intersection of the femoral mechanical axis and tibial mechanical axis [3]. Varus and valgus deformities are common clinical findings in knee OA and there is evidence that coronal malalignment may be a factor in the development and progression of the disease [4,5]. The extent to which any deformity in the coronal plane is correctable on examination may influence the surgical technique used in TKA. There is currently no accepted means of quantifying this reliably, and a non-invasive method of doing could be a valuable tool in operative planning.

[Figure 1 here]

The currently accepted best surgical practice is to place femoral and tibial TKA components such

that the post-operative mechanical axis of the lower limb is  $0^\circ$ , that is, a straight line from hip to knee to ankle centres. The standing, antero-posterior (AP) long-leg radiograph is a reliable method of measuring mechanical alignment in the clinical setting [6,7]. A well-described aim of TKA is to obtain a post-operative MFTA within a margin of error of  $0\pm 3^\circ$ . These parameters have been advocated by a number of clinical studies, with TKAs outlying  $3^\circ$  of a 'neutral' alignment of  $0^\circ$  associated with decreased survivorship [8,9]. This has been corroborated by retrieval studies which suggest that excessive post-operative varus or valgus alignment can lead to increased implant wear [10-12].

Evidence for the significance of MFTA was one of the key rationales for the development of knee navigation, a form of computer-assisted orthopaedic surgery (CAOS). The goal of navigated knee arthroplasty is to increase the precision of implant positioning - to place components more frequently within the biomechanical parameters thought to be associated with the highest patient function and greatest survivorship. A number of randomised control trials [13-15] and meta-analyses [16,17] have suggested that knee navigation systems can improve precision of implant alignment within the margin of error of  $0\pm 3^\circ$ , which may lead to improvements in functional scores and survivorship. However, the belief that the 'well-aligned' TKA equates to improved outcome has been challenged by some authors [18,19]. Knee navigation has also been used in the intra-operative process of soft-tissue balancing, which aims to provide an arthroplasty which exhibits coronal stability through the range of movement [20,21].

Mechanical alignment of the lower limb is thus relevant to navigated TKA on a number of levels. A non-invasive navigation system providing a reliable measure of MFTA through the range of knee flexion could be of clinical value in assessing knee deformity and laxity pre-operatively, and alignment and balancing post-operatively, with the distinct advantage of avoiding ionizing radiation.

Our research uses a non-invasive version of an image-free navigation system. Image-free navigation does not require pre or intra-operative radiological imaging, relying instead on localisation of anatomical landmarks with infrared emission. Infrared trackers are fixed to bony anatomical landmarks by securing to drilled cortical bone screws, which facilitates the measurement of knee alignment and kinematics intra-operatively. The non-invasive system used in this study was initially developed by Clarke *et al* [22] using navigation software originally designed for high tibial osteotomies. In this system, passive infrared trackers were secured to the skin surface using elastic strapping. In a group of 30 volunteers, with the knee in full extension (supine and with bipedal stance) an intra-observer repeatability of Coefficient of Repeatability (CR) and Limits of Agreement (LOA) of <2 were found [22].

More recently, software has been specifically developed for non-invasive navigation, as used in this study (Physiopilot™ v1.0; B. Braun Aesculap, Tuttlingen, Germany). In a cadaveric study, Russell *et al* [23] found an acceptable intra-observer repeatability ( $CR \leq 2$ ) for measurement of MFTA with the Physiopilot™ system through the range of knee flexion 0-60° [23]. Our research is the first in-vivo study of this novel software, which is designed specifically for non-invasive use.

## **Materials and Methods**

The aim of our study was to evaluate a non-invasive knee navigation system in the measurement of lower limb mechanical femoro-tibial angle (MFTA), with manually applied varus and valgus force, in knee extension and through the range of knee flexion. The non-invasive system consisted of an infrared camera, navigation base-plates, fabric strapping, passive infrared trackers, navigation pointer, and the Physiopilot® v.1.0 software package (all B. Braun Aesculap, Tuttlingen, Germany). An overview of the system setup is illustrated in Figure 2.

[Figure 2 here]

The system used was a non-invasive, image-free knee navigation system similar to that originally developed by Clarke *et al* [22] and Russell *et al* [23]. It aims to provide reliable kinematic data for the tested knee, including real-time mechanical alignment, similar to the measurements which would be obtained by using a conventional invasive image-free navigation package in the operating theatre. We aimed to perform varus and valgus stress testing through the range of 0°-90° with 15° increments, to give an experiment protocol which would be practical yet also reflect the intra-operative process of soft-tissue balancing [24] and allow identification of trends in coronal laxity with increasing flexion. There were two operators of the system, who were the authors of this study.

Passive infrared trackers were mounted to base-plates, as illustrated in Figure 3. These were originally designed for use in theatre to secure navigation trackers to the dorsum of the foot in the registration of the kinematic centre of the ankle. Base-plates were secured to the skin surface using fabric strapping. This consisted of 45mm width elastic webbing with a series of eyelets for attachment of the base-plate. Various lengths of strapping were available, allowing the base-plate to be secured firmly to differing thigh and calf diameters.

A group of 23 participants was recruited, all with asymptomatic, healthy knees; 13 male and 10 female, with mean age 32.6 years (range 23-59 years) and mean BMI 24.1 (range 18.8-43.1). This gave an initial sample of 46 knees. Exclusion criteria were previous knee surgery, known knee pathology, and skin lesions, oedema or poor skin quality which would have risked cutaneous damage when attaching the infrared trackers. University Departmental Ethics Committee approval was granted for testing the Physiopilot™ system and all volunteers signed a written informed consent proforma.

[Figure 3 here]

The participant was asked to lie supine on an examination trolley. Base-plates with trackers were attached to the lower limb, with positions based on manufacturer's recommendations for tracker placement in use of the conventional, invasive Orthopilot® navigation system (B. Braun Aesculap, Tuttlingen, Germany), as illustrated in Figure 3. The thigh tracker, corresponding to the recommended position of a distal femoral navigation tracker, was secured to the skin overlying the vastus medialis anteriorly, just proximal to the superior pole of the patella. The leg tracker, corresponding to the recommended position of a tibial navigation tracker, was secured to the skin overlying the tibial crest, anteriorly, at the level of mid-leg.

The protocol followed the standardised workflow for registration as per the Physiopilot™ software package. Palpable bony landmarks were registered with the navigation pointer, in turn; medial epicondyle, lateral epicondyle, centre of the knee joint (Figure 4), medial malleolus, lateral malleolus, centre of the ankle joint (Figure 5). With the limb relaxed and the knee extended, and avoiding movement of the pelvis, the centre of the hip joint was registered by the operator holding the heel and passively circumducting the lower limb. The Physiopilot™ registration process was then completed by passively flexing the knee from the limit of extension to 90° of flexion, and internally and externally rotating the tibia to end points at 90° of flexion.

Following the registration process, with participant supine and knee in extension, the MFTA calculated by the software was displayed on-screen and recorded.

- With the knee extended in an 'unlocked' position, a varus force was applied manually by the operator until an end-point in the knee joint was reached. A valgus force was then applied manually by the operator until end-point was reached. The variation/range of MFTA with

applied force was displayed on-screen and recorded.

- Force was applied with one hand on either the medial or lateral side of the distal femur to stabilise the thigh, and the other hand gripping the participant's ankle, to stress the tibia (and hence the knee joint itself) in a varus or valgus direction in the coronal plane.
- The knee was then tested through the range of flexion to 90° in 15° increments; with the on-screen flexion angle at 15°, 30°, 45°, 60°, 75°, 90°, a varus and then valgus force was applied manually by the operator until an end-point was reached. The range/variation of MFTA with applied force was recorded.
- The Physiopilot™ workflow was completed by recording the bipedal and monopedal MFTA; for the latter the participant was asked to balance on the tested foot.

[Figure 4 here]

[Figure 5 here]

The trackers, strapping and base-plates were then removed, and the protocol was repeated on the same knee, this time by the second operator. Each operator performed two registrations on each knee. Each participant therefore underwent eight registrations with the Physiopilot™ software, four on each knee.

In order to avoid the operator obscuring the field of the infrared camera, and to avoid the operator having to lean over the contralateral limb to conduct stress testing (which may have led to alteration in technique), the patient was asked to turn on the examination trolley when switching between testing of the right and left knees. As can be seen in Figure 3, the lower limb contralateral to the tested knee is closer to the infrared camera, and we found this patient positioning facilitated easier, and more reproducible, stress testing of the knee.



Appropriate intra-observer and inter-observer statistics were calculated, which gave repeatability for a single user and inter-observer correlation. Additional statistical tests were carried out to determine any effect of participant sex, BMI or age on observed coronal laxities.

## **Results**

The efficacy of the Physiopilot™ system as a measurement tool for mechanical alignment was evaluated by calculating intra-observer Coefficients of Repeatability (CR) and inter-observer Intraclass Correlation Coefficients (ICC).

Using a 95% confidence interval, CR gives a value below which the absolute difference between consecutive measurements would occur with a probability of 0.95 [26]. The lower the CR, the less likely there is to be a large discrepancy between repeated measurements taken by a user and thus the more reliable or ‘repeatable’ the tool. CR is an established means of assessing measurement tools where repeated measurements of a variable are taken by a user. Values for each operator were calculated using the standard deviation of the absolute difference between the first and second readings taken for each knee, where  $CR = 1.96 \times \sqrt{\Sigma(d2-d1)^2/(n-1)}$ . This statistic accounts for both random and systematic errors in measurement [26]. While conventional, clinically approved knee navigation systems aim for a maximum error of 1° [27], for the non-invasive system a maximum CR of 2 was considered as acceptable due to the likelihood of additional errors incurred by soft tissue artefacts [22].

Inter-observer correlation was assessed with calculation of ICC, an accepted method in clinical trials of measuring correlation between two methods of measurement when taking repeated

measurements [28]. Each operator (A and B) performed two registrations of the Physiopilot™ system on each knee, with each user's registration taken as a representation of an individual 'method' of examining the knee for the purposes of this model. Using the two readings taken by each operator, a mean was calculated and used to find ICC [29]. Calculated as a value between 0 and 1.00, a higher value of ICC indicates a greater degree of correlation between two users' measurements of the same variable. Correlation is considered 'poor' when  $ICC < 0.40$ , fair with  $ICC 0.40-0.70$ , and good when  $ICC > 0.70$  [30]. ICC was calculated using SPSS statistics software (SPSS v.20.0; IBM Corp, Armonk, NY, USA).

Intra-observer repeatabilities are illustrated in the form of Bland-Altman plots. The mean of each user's two readings is plotted against the difference between the first and second readings. Points which are duplicate appear as bold. Using a 95% confidence interval, Limits of Agreement (LOA), indicated by the blue lines, give the values with which the difference between the first and second reading would lie with a probability of 0.95, where  $LOA = \text{mean difference} \pm 1.96 \times \sqrt{\Sigma(d2d1)^2/(n-1)}$  [28]:

### ***Limits of agreement in extension***

Analysis of data from registrations for our first 8 volunteers (equating to 16 knees) found that there was inadequate initial agreement for supine MFTA in extension. As with the trial of non-invasive navigation conducted by Clarke et al [22], repetitions of the registration process which did not give an initial measurement of MFTA in extension within  $\pm 2^\circ$  of the first registration were discounted, and the registration workflow was restarted. In the use of invasive navigation systems in the operating theatre, it also is common practice for initial registrations of MFTA in extension which are thought to be erroneous by the surgeon to be discounted and the registration process repeated [31]. A difference in agreement of greater than  $2^\circ$  in extension MFTA was likely to lead to unsatisfactory agreement for MFTA in flexion. Data from the first 8 consecutive participants was

therefore excluded due to unsatisfactory agreement between repeated registrations - that is, a difference of  $>2^\circ$  between each operator's 1<sup>st</sup> and 2<sup>nd</sup> registrations. This gave a final sample of 15 volunteers, equating to 30 knees. Results are illustrated in Figures 6 and 7, which show satisfactory agreement of supine MFTA in extension.

[Figure 6 here]

[Figure 7 here]

### ***Agreement with applied coronal force***

A varus and valgus force was applied manually to the knee to the palpable limit of coronal laxity. These manoeuvres were initially performed in extension with the knee in an 'unlocked' position. Figures 8-11 show that, for both users, intra-observer repeatability falls outwith acceptable limits for measurement of MFTA when force is applied.

[Figure 8 here]

[Figure 9 here]

[Figure 10 here]

[Figure 11 here]

Figures 12-15 illustrate intra-observer agreement at  $90^\circ$  of flexion, which also fall outwith acceptable limits.

[Figure 12 here]

[Figure 13 here]

[Figure 14 here]

[Figure 15 here]

### ***MFTA in extension and mean coronal laxity***

Table 1 shows the mean MFTA in extension and stance found for each user, and Figures 16 and 17 illustrate the mean varus and valgus knee angulations found for each user as an illustration of the trend found in coronal range of movement. The given angles are the on-screen MFTAs displayed with the knee taken to its palpable limits of coronal laxity at the measured increments of flexion.

[Table 1 here]

[Figure 16 here]

[Figure 17 here]

The overall trends for Figures 16 and 17 are similar, with an increase in coronal laxity as flexion increases - that is, with increasing increments of flexion, both users found that a greater increase in MFTA in a varus or valgus angulation was measured by the non-invasive system. In addition, the general trend in laxity illustrated is for greater varus angulation than valgus when force is applied. These results are in keeping with recent studies of coronal laxity [32,33].

### ***Intra-observer repeatability and inter-observer correlation***

Tables 2-5 summarise the intra-observer CRs and inter-observer ICCs, There is satisfactory intra-observer agreement in extension, however this remains unacceptable with applied force through the range of flexion. The general trend is again towards increasing loss of agreement with increasing flexion, especially with varus force. Inter-observer correlation remains fair throughout.

[Table 2 here]

[Table 3 here]

[Table 4 here]

[Table 5 here]

### ***Analysis of sex, BMI and age***

Further statistical testing was carried out to determine if the independent variables of sex, BMI and age had any statistically significant effect on the outcomes of MFTA in extension and coronal laxity in extension and flexion with applied force. In this analysis non-parametric tests were selected such that there was no assumption that the measured outcomes followed a normal distribution. The level of significance was taken as 0.05. For the ordinal variable of sex, a Mann-Whitney U-test was conducted to determine if there were any significant differences in MFTA and laxity for either user at the increments of flexion. BMI and age were taken as continuous variables and therefore a Spearman's rank correlation coefficient was calculated, with values closer to 1.00 indicating a greater degree of correlation [34]. All testing was performed with SPSS statistics software package (SPSS v.20.0; IBM Corp, Armonk, NY, USA). P-values correct to 3dp at extension and each increment of flexion are given in Tables 6-8.

[Table 6 here]

[Table 7 here]

[Table 8 here]

## **Discussion**

A number of studies have suggested that MFTA following TKA (with or without use of a navigation system) can have a significant effect on patient functional outcome and implant survivorship [8,9,35,36]. We therefore felt that there was a justifiable rationale for a study of coronal alignment in the use of the non-invasive navigation system. With values of 1.55 and 1.33 calculated for intra-observer Coefficients of Repeatability (CR), and 0.72 calculated for inter-observer Intraclass Correlation Coefficient (ICC), our study found that after each user trialling 8 participants, an acceptable level of repeatability for the measurement of supine mechanical alignment in extension was achieved. These results are in keeping with the findings previously demonstrated by Russell *et al* [23] and Clarke *et al* [22]. Our view is that initial use of the system represented a ‘learning-curve’ of registrations required before satisfactory, repeatable use of the system is achieved. There is evidence in the literature that more precise results are obtained with navigation systems after an initial ‘training’ period of usage - as yet not clearly quantified.

Industry standards for operative CAOS systems state that navigation should aim for an error of 1° or less in measurement of alignment, while a margin of error of  $\pm 3^\circ$  from neutral is frequently accepted in the literature as an aim for post-operative lower limb alignment. While our calculations show fair inter-observer ICC values for applied force through the range of flexion, CR values show inadequate intra-observer agreement when varus or valgus force is applied, both in extension and

through the range of flexion from 0-90°. In addition, CR and ICC fall outwith acceptable limits for extension MFTA with bipedal and monopodal stance. The general trend is for fair inter-observer correlation in flexion, which does suggest a degree of similarity between the MFTAs measured by each user at each increment of flexion.

Our study of the Physiopilot™ system at this point in time did not involve radiographic assessment of volunteers - for example standing long-leg radiographs. Volunteers did not undergo TKA subsequently, and thus did not have alignment measured by a conventional, invasive knee navigation system. Our research therefore does not assess agreement of the Physiopilot™ system measurements with a ground truth or gold-standard measurement, but rather evaluates the system in terms of intra-observer and inter-observer agreement. Despite this, in keeping with previous research, our work can support the use of the Physiopilot™ system in the measurement of lower limb MFTA in extension. However there is insufficient evidence from our trial to support its use in measuring MFTA with manually applied force in standard clinical examination.

We believe that there were a number of sources of error which could account for unsatisfactory reliability of the non-invasive system with applied force in flexion.

### ***Soft tissue artefacts***

Conventional, invasive image-free knee navigation systems use passive or active infrared trackers which are attached to bony anatomy with drilled cortical bone screws. They are assumed to be fixed and immune - as possible - to any artefactual movement which would result in loss of registered position. It is therefore intuitive to assume that attaching trackers non-invasively to the skin surface could result in a significant level of error.

Attempts were made to minimise these sources of error by careful prior definition of the surface

anatomical landmarks to be registered in each trial participant. In addition, in performing the examination on each patient we aimed to direct varus or valgus force only through the medial or lateral malleolus while steadying the proximal thigh; if soft tissue movement was seen in the calf or thigh then the manoeuvre was repeated.

In performing varus and valgus stress testing, force was applied directly to the skin, subcutaneous tissue and muscle in close proximity to the attached markers. It is therefore assumed that errors in marker position would be incurred due to deformation and movement of these structures. Reflective infrared markers used in the Physiopilot™ system are attached to metal base plates which project the markers upwards. Movement of underlying muscle and tissue may lead to a higher relative movement in positioning of markers due to a pendulous movement of the cluster set moving to either side on its mounting. Additionally, conventional invasive navigation systems are used in anaesthetised patients with muscle relaxation, where variation in limb muscle activation and tone is negligible. While patients in the operating theatre are supine, our volunteers were assessed with bipedal and monopodal stance. It is possible that these factors may all lead to error in registered position of markers attached to the skin.

In initial studies by Clarke *et al* [22], attempts were made to quantify the errors incurred by using a navigation system non-invasively. Varus and valgus stress testing of the knee was found to increase the limits of agreement by  $\pm 0.5$  when supine MFTA in extension was repeated. Bipedal stance was associated with an  $\pm 0.2$  increase in limits of agreement [22]. These calculations give an estimate of the error which may be incurred by muscle activation and weight-bearing in stance, and with muscle activation due to stress testing. However, it is important to note that these quantifications relate specifically to the repeat measurement of supine extension MFTA after stress testing or weight-bearing, which cannot be easily extrapolated to the stress testing in flexion performed in our trial. In addition, the studies by Clarke *et al* found a CR of 3 for measurement of MFTA with



bipedal stance, which indicates a loss of repeatability for weight-bearing versus measurements when supine [22]. Our own values for this were higher; 4.02 and 6.43 for Users A and B, respectively. It is possible that increased muscle activation while standing leads to a soft tissue artefact and change in positioning of skin surface markers which incurs significant error.

A number of other studies have suggested that there may be significant errors associated with soft tissue movement in using optical trackers. Sati *et al* [37] found that skin overlying the medial and lateral condyles could move in position by up to 17mm using trackers, while Stagni *et al* [38] found standard deviations of movement of skin surface trackers from original positions of 31mm for the thigh and 21mm for the calf when limb movement was introduced. Manal *et al* [39] compared skin-mounted trackers with trackers fixed to bony landmarks in gait analysis, finding an error of 1 Nm was found in moment about a longitudinal axis.

In addition, applying coronal force with the introduction of knee flexion invariably results in a rotational movement of the femur and hip joint. In our study the examination technique attempted to minimise lower limb rotation in assessing extension MFTA and coronal laxity with one hand firmly holding the thigh/femur while the other hand applied force from the ankle. However it was impossible to eliminate rotation of the tested limb, particularly when flexion was introduced. Internal and external rotation of the femur with knee flexion and extension have been well described in knee biomechanics [40], although the potential effect of this on the measurement of coronal alignment by a navigation system has not. Mayr *et al* [41] found that flexion of the hip joint to 90° could result in an error of up to 2.5° in the mechanical alignment measured using a corresponding distal femoral navigation tracker. Kannan *et al* [42] measured mechanical alignment in a prosthetic model of the lower limb using long-leg radiographs with 5° increments of knee flexion and lower limb internal and external rotation up to 20°. A combination of knee flexion and lower limb external rotation was found to progressively alter measured MFTA up to an error of 5°.

Studies of this nature suggest that if rotation is introduced to the lower limb when flexing the knee, a significant error can occur in assessment of MFTA

### ***Learning-curve in CAOS***

While the operators conducted multiple registrations (>20) on each other as test subjects before commencing the trial, it was felt that relative inexperience with knee navigation systems was a significant source of error and contributed to the unsatisfactory values for repeatability with flexion. There is evidence in the literature to suggest that there is a significant learning curve in the use of navigation systems before reliable, repeatable results are produced [43,44] which one study of navigated TKA quantified as 20 procedures [31]. We feel that this is evident in our work, in that repeated registrations in our initial volunteers found unsatisfactory agreement in extension. There was, however, satisfactory initial agreement after each operator had performed 32 registrations in the trial.

Difficulties with adapting clinical examination technique with use of the non-invasive system hardware may have led to error when flexion was introduced. This may be particularly related to user inexperience with navigation systems. In stress testing participants, care was taken by each user to perform stress manoeuvres which were as reproducible as possible, with one hand placed on the medial/lateral malleolus and medial/lateral knee joint margin to apply a coronal force. However, it was essential not to accidentally move the marker sets attached to the skin surface in doing so, or to obscure the markers from the view of the infrared camera. This inherently made clinical examination a more difficult process.

### ***Force standardisation***

As would be performed in a routine clinical examination of the knee, the joint was stressed with varus and valgus force to what was felt to be an 'end-point' - as deemed subjectively by the

operator. There was therefore no quantification of the force (torque) applied to the joint.

Measurement of applied torque is not viewed as a standard aspect of clinical examination of the knee, and there is no consensus found in the orthopaedic literature as to a preferred method of doing so.

‘Force application devices’ have previously been used to standardise coronal force applied to the knee, such as hand-held brackets fitted with torque sensors [46]. It was not technically possible at the time of our study to incorporate such hardware into the Physiopilot™ workflow. In a previous, cadaveric study of the Physiopilot™ system, torque in the coronal plane was standardised at 15Nm using a manual hand held force transducer. This equipment was attached to the distal tibia using cortical bone screws, however, which was not suitable for our in-vivo trial [23]

### ***Effects of sex, BMI and age***

As illustrated in Figures 16-17, varus and valgus laxity was seen to increase with increasing flexion. However, the repeatability of the non-invasive system was also seen to decrease with increasing flexion; one may therefore envisage that if a participant factor (for example age) were to affect coronal laxity, this may in turn affect the precision of the system.

There was a tendency observed for a statistically significant difference between males and females in coronal laxity with increased knee flexion, particularly at 60° and greater. Increased generalised ligamentous laxity in female knees has been described in a number of studies, particularly in relation to AP laxity with the ACL and PCL [47]. This has also been observed in the MCL and LCL [48]. It would therefore not be unexpected to encounter an effect with sex in our results, however given the poor precision of the Physiopilot™ system with increasing flexion it would be difficult to consider the results as convincing evidence of sex as a significant factor in coronal laxity.

As seen in Table 7, Spearman rank correlation coefficients for BMI have largely (although not universally) negative values for both users at increments  $> 0^\circ$ , suggesting a tendency for increased BMI to decrease laxity as flexion is introduced to the tested knee. However, only two of these values correspond to a statistically significant effect at the 0.05 level, and indeed user B found a significant positive correlation for BMI with varus force in extension. It is therefore difficult to reach any definitive conclusion on the effect of BMI on our use of the non-invasive system.

As illustrated in Table 8, the largely negative values for Spearman rank correlation coefficients for increasing age with coronal laxity in flexion suggests that laxity would tend to decrease with age, however this again is not uniform for both users and only two values are at a level of statistical significance. There is limited evidence of a relationship between age and coronal knee alignment and laxity [49].

## **Conclusion**

The clear issue with the Physiopilot™ non-invasive system which arises from our study is the finding of unsatisfactory reliability for coronal alignment in flexion, and unsatisfactory reliability when varus or valgus force is applied, both in extension and through the range of flexion 0-90°. With potential sources of error including soft tissue and skin deformation, muscle activation and femoral rotation, it is our opinion that further non-clinical studies (cadaveric or in-vivo) are required to address these issues. Nevertheless, the Physiopilot system was able to give an acceptably repeatable measure of static MFTA in extension when supine, and a non-invasive image-free system has the distinct advantage of avoiding the ionizing radiation incurred by long-leg radiographs or CT scanning. The potential to conduct further studies of coronal alignment in larger samples of both healthy and osteoarthritic patients may develop a better understanding of ‘normal’ MFTA versus alignment in OA, and guide a standardised/algorithmic approach to surgical technique to establish

the most appropriate target alignment in TKA. Monitoring of coronal alignment in long-term follow-up post-arthroplasty may also help to determine its significance in functional outcome and survivorship.

### **Declaration of interest statement**

No grants or funding were received in order to conduct the research presented in this paper. The authors do not hold any positions with any manufacturers which could constitute a conflict of interest.

### **Word Count**

### **References**

- [1] Felson DT. An update on the pathogenesis and epidemiology of osteoarthritis. Radiologic Clinics of North America, 42(1)(2004), pp.1-9
- [2] Powers-Freeling L (Chair). The 11<sup>th</sup> Annual Report of the National Joint Registry for England, Wales and Northern Ireland.
- [3] Cooke TDV, Sled EA, Scudamore RA. Frontal plane knee alignment: A call for standardized measurement. Journal of Rheumatology 34(9)(pp 1796-1801),2007
- [4] Sharma L, Song J, Felson DT et al. The role of knee alignment in disease progression and functional decline in knee osteoarthritis. JAMA 286(2):188-95,2001 Jul 11

- [5] Cerejo R, Dunlop DD, Cahue S et al. The influence of alignment on risk of knee osteoarthritis progression according to baseline stage of disease. *Arthritis & Rheumatism* 46(10):2632-6,2002 Oct
- [6] Babazadeh S, Dowsey MM, Bingham RJ et al. The long leg radiograph is a reliable method of assessing alignment when compared to computer-assisted navigation and computer tomography. *Knee* 20(4):242-9,2013 Aug
- [7] Rauh MA, Boyle J, Mihalko WM et al. Reliability of measuring long-standing lower extremity radiographs. *Orthopedics*, 30(2007),pp.299-303
- [8] Jeffrey RS, Morris RW, Denham RA. Coronal alignment after total knee replacement. *Journal of Bone & Joint Surgery - British Volume* 73(5):709-14,1991 Sep
- [9] Ritter MA, Faris PM, Keating EM et al. Post-operative alignment of total knee replacement: its effect on survival. *Clinical Orthopaedics*;299:153-156,1994
- [10] Wasielewski RC, Galante JO, Leighty RM et al. Wear patterns on retrieved polyethylene tibial inserts and their relationship to technical considerations during total knee arthroplasty. *Clinical Orthopaedics & Related Research*. (299):31-43,1994 Feb
- [11] Collier MB, Engh CA Jr, McAuley JP et al. Factors associated with the loss of thickness of polyethylene tibial bearings after knee arthroplasty. *Journal of Bone & Joint Surgery - American Volume* 89(6):1306-14,2007 Jun
- [12] Werner FW, Ayers DC, Maletsky LP et al. The effect of valgus/varus malalignment on load distribution in total knee replacements. *Journal of Biomechanics* 38(2):349-55,2005 Feb
- [13] Huang NF, Dowsey MM, Ee E et al. Coronal alignment correlates with outcome after total knee arthroplasty: five-year follow-up of a randomized controlled trial. *Journal of Arthroplasty* 27(9):1737-41,2012 Oct
- [14] Cheng T, Zhang G, Zhang X. Clinical and radiographic outcomes of image-based computer-assisted total knee arthroplasty: an evidence-based evaluation. *Surg Innov* 2011 Mar;18(1):15-20

- [15] Choong PF, Dowsey MM, Stoney JD. Does accurate anatomical alignment result in better function and quality of life? Comparing conventional and computer-assisted total knee arthroplasty. *Journal of Arthroplasty* 24(4):560-9,2009 Jun
- [16] Rebal BA, Babatunde OM, Lee JH et al. Imageless Computer Navigation in Total Knee Arthroplasty Provides Superior Short Term Functional Outcomes: A Meta-Analysis. *The Journal of Arthroplasty* Volume 29, Issue 5, May 2014, Pages 938–944
- [17] De Steiger RN, Liu YL, Graves SE. Computer Navigation for Total Knee Arthroplasty Reduces Revision Rate for Patients Less Than Sixty-five Years of Age. *Journal of Bone and Joint Surgery - American* Volume 2015 Apr 15;97(8):635-642
- [18] Hiscox CM, Bohm ER, Turgeon TR et al. Randomized trial of computer-assisted kneearthroplasty: impact on clinical and radiographic outcomes. *Journal of Arthroplasty* 26(8):1259-64, 2011 Dec
- [19] Parratte S, Pagnano MW, Trousdale RT et al. Effect of Postoperative Mechanical Axis Alignment on the Fifteen-Year Survival of Modern, Cemented Total Knee Replacements. *J Bone Joint Surg Am* 2010 Sep 15;92(12):2143-9
- [20] Hakki S, Coleman S, Saleh K et al. Navigational predictors in determining the necessity for collateral ligament release in total knee replacement. *Journal of Bone & Joint Surgery - British* Volume 91(9):1178-82,2009 Sep
- [21] Picard F, Deakin AH, Clarke IV et al. A quantitative method of effective soft tissue management for varus knees in total knee replacement surgery using navigational techniques. *Proceedings of the Institution of Mechanical Engineers Part H - Journal of Engineering in Medicine* 221(7):763-72,2007 Oct
- [22] Clarke JV, Riches PE, Picard F, Deakin AH. Non-invasive computer-assisted measurement of knee alignment. *Comput Aided Surg* 2012;17(1):29-39
- [23] Russell D, Deakin A, Fogg QA et al. Non-invasive quantification of lower limb mechanical alignment in flexion. *Comput Aided Surg* 2014;19(4-6):64-70

- [24] Insall JN. Correction of arthritic deformities of the knee. In: McCarty DJ, ed. Arthritis and allied conditions: a textbook of rheumatology. Tenth edition. Philadelphia: Lea and Febiger, 1985:771-84
- [25] Rajadhyaksha AD, Mehta H, Zelicof SB. Use of tibialis anterior tendon as distal landmark for extramedullary tibial alignment in total knee arthroplasty: an anatomical study. American Journal of Orthopaedics 2009 Mar;38(3):E68-70
- [26] Vaz S, Falkmer T, Passmore AE et al. The Case for Using the Repeatability Coefficient When Calculating Test-Retest Reliability. PLoS One 2013;8(9):e73990J
- [27] American Society for Testing and Materials. ASTM F2554-10: Standard Practice for Measurement of Positional Accuracy of Computer Assisted Surgical Systems. 2010: Pennsylvania; ASTM International
- [28] Bland JM, Altman DG. Statistical methods for assessing agreement between two methods of clinical measurement. Lancet 1986 Feb 8;1(8476):307-10
- [29] McGraw KO, Wong SP. Forming inferences about some intraclass correlation coefficients. Psychological Methods Vol 1(1) Mar 1996 30-46
- [30] Fleiss JL. The Design and Analysis of Clinical Experiments. Wiley:New York;1986
- [31] Jenny JY, Miehke RK, Giurea A. Learning curve in navigated total knee replacement. A multi-centre study comparing experienced beginner centres. Knee 2008 Mar;15(2):80-4
- [32] Okazaki K, Miura H, Matsuda S et al. Asymmetry of mediolateral laxity of the normal knee. J Orthop Sci 2006 May;11(3):264-6
- [33] Heesterbeek PJ, Verdonschot N, Wymenga AB. In vivo knee laxity in flexion and extension: a radiographic study in 30 older healthy subjects. Knee 2008 Jan;15(1):45-9
- [34] Corder GW, Foreman DI. Nonparametric statistics for non-statisticians: a step-by-step approach. 2009;New Jersey:Wiley & Sons
- [35] Berend ME, Ritter MA, Meding JB et al. Tibial component failure mechanisms in total knee arthroplasty. Clinical Orthopaedics & Related Research (428):26-34,2004 Nov



- [36] Kamat YD, K.M. Aurakzai, A.R. Adhikari et al. Does computer navigation in total knee arthroplasty improve patient outcome at midterm follow-up? *Int Orthop*, 33(2009), p.1567
- [37] Sati M, de Guise JA, Larouche S et al. Quantitative assessment of skin-bone movement at the knee. *The Knee* 3(1996)pp121-138
- [38] Stagni R, Fantozzi S, Cappello A et al. Quantification of soft tissue artefact in motion analysis by combining 3D fluoroscopy and stereophotogrammetry: a study on two subjects. *Clinical Biomechanics* 20(3):320-9,2005 Mar
- [39] Manal K, McClay I, Richards J et al. Knee moment profiles during walking: errors due to soft tissue movement of the shank and the influence of the reference coordinate system. *Gait Posture* 2002 Feb;15(1):10-7
- [40] Johal P, Williams A, Wragg P et al. Tibio-femoral movement in the living knee. A study of weight bearing and non-weight bearing knee kinematics using 'interventional' MRI. *J Biomech.* 2005 Feb;38(2):269-76
- [41] Mayr E, Moctezuma de la Barrera JL, Eller G, et al. The effect of fixation and location on the stability of the markers in navigated total hip arthroplasty: a cadaver study. *J Bone Joint Surg [Br];88-B:168-72,2006*
- [42] Kannan A, Hawdon G, McMahon SJ. Effect of flexion and rotation on measures of coronal alignment after TKA. *J Knee Surg* 2012 Nov;25(5):407-10
- [43] Donnelly WG, Crawford RW, Whitehouse SL et al. Knee Navigation: The Learning Curve. *J Bone Joint Surg Br* 2004 vol 86-B Supp. IV 474-475
- [44] Maniar RN, Johorey AC, Pujary CT et al. Margin of error in alignment: a study undertaken when converting from conventional to computer-assisted total knee arthroplasty. *Journal of Arthroplasty* 26(1):82-7,2011 Jan
- [45] Unitt L, Sambatakakis A, Johnstone D et al. Short-term outcome in total knee replacement after soft-tissue release and balancing. *Bone Joint Surg Br* 2008 Feb;90(2):159-65
- [46] Clarke JV, Wilson WT, Wearing SC et al. Standardising the clinical assessment of coronal

knee laxity. Proc Inst Mech Eng H 2012 Sep;226(9):699-708

[47] Rozzi SL, Lephart SM, Gear WS et al. Knee joint laxity and neuromuscular characteristics of male and female soccer and basketball players. Am J Sports Med. 1999 May-Jun;27(3):312-9

[48] Shultz SJ, Schmitz RJ, Beynnon BD. Variations in Varus/Valgus and Internal/External Rotational Knee Laxity and Stiffness across the Menstrual Cycle. J Orthop Res. 2011 Mar; 29(3):318-325

[49] Sharma L, Lou C, Felson DT et al. Laxity in healthy and osteoarthritic knees. Arthritis & Rheumatism 42(5):861-70, 1999 May



Figure 1: Illustration of relevant anatomical and mechanical lower limb axes. The mechanical femoro-tibial angle (MFTA) is found between the red and blue lines



Figure 2: Physiopilot™ system set-up. The camera, positioned on the tripod, acts as the source of infrared emission and detection. Trial participants are positioned, initially supine, on the examination trolley and stand in front of the trolley for bipedal and monopedal stance. To the right of the picture the Physiopilot™ workflow is shown on screen.

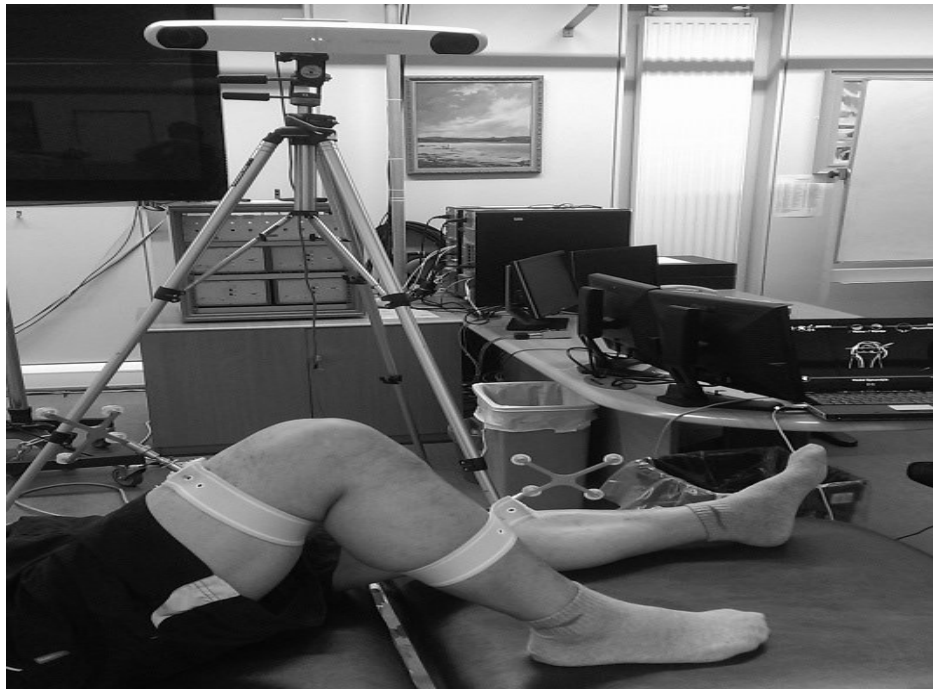


Figure 3: Passive navigation tracker clusters are secured to the thigh and leg with fabric strapping. The examination trolley is positioned so that the tested limb is approximately 2m from the camera, as per manufacturer's recommendations.



Figure 4: Registration of knee centre. The pointer is placed anteriorly, in the pre-patellar area, aiming towards the tibial spines. This corresponded to the midway point on screen between the registered medial and lateral epicondyles.



Figure 5: Registration of ankle centre. The pointer was aimed at the lateral border of the palpable tibialis anterior tendon, which has been shown to be a reproducible landmark clinically in registration of the kinematic centre of the ankle joint [25]. On screen, we aimed for this to correspond to midway between the registered medial and lateral malleoli.

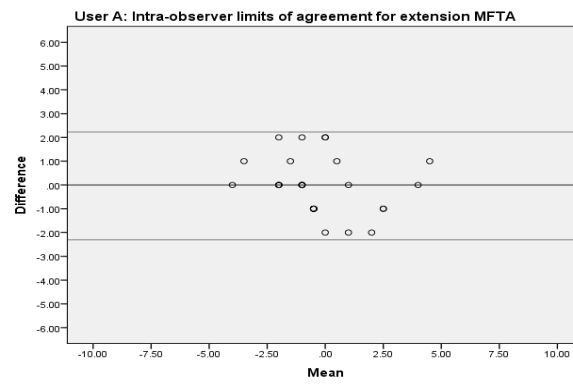


Figure 6: Bland-Altman plot for intra-observer repeatability for User A in measurement of supine MFTA with knee in extension.



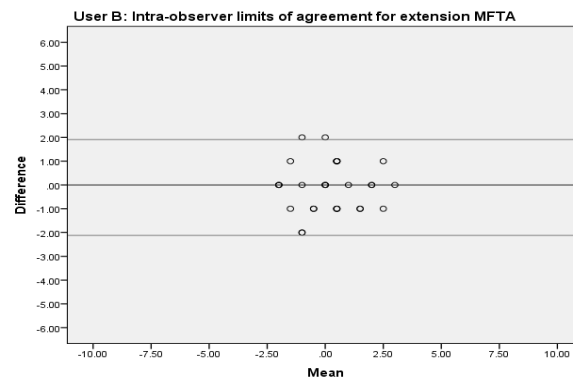


Figure 7: There is satisfactory agreement for User B in supine MFTA in extension.

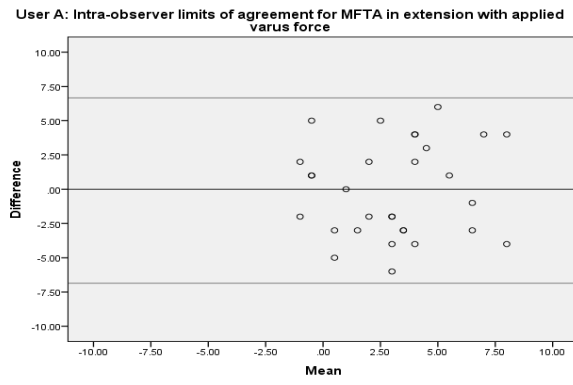


Figure 8: Limits of Agreement are greater than the accepted values when a varus force is applied.

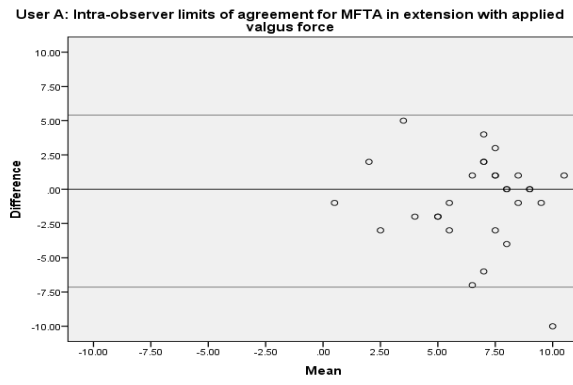


Figure 9: Bland-Altman plot illustrating intra-observer agreement for MFTA in extension with valgus force for User A.

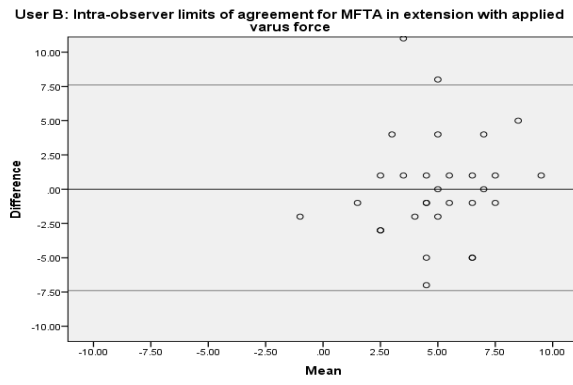


Figure 10: Bland-Altman plot illustrating intra-observer agreement for MFTA in extension with varus force for User B.

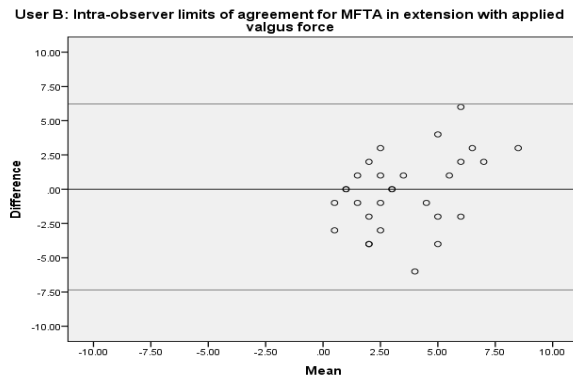


Figure 11: Bland-Altman plot illustrating intra-observer agreement for MFTA in extension with valgus force for User B.

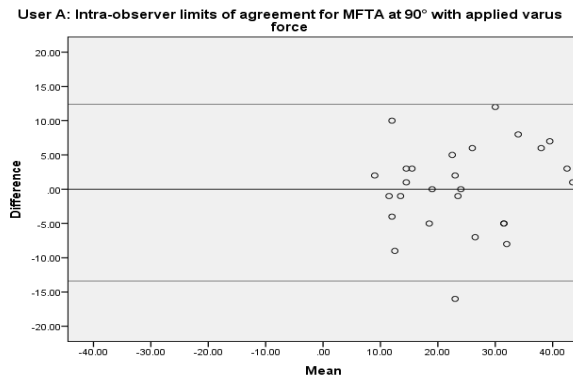


Figure 12: There is a significant increase in the Limits of Agreement for measured MFTA at 90° flexion with applied force.

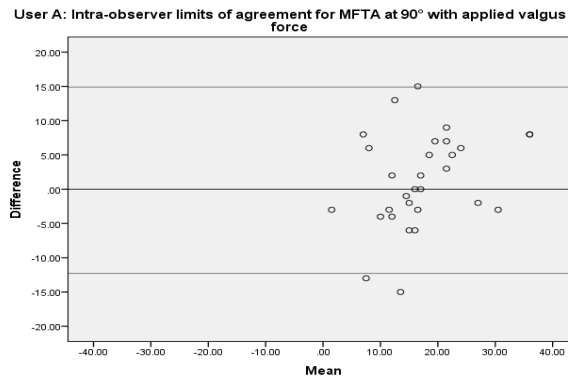


Figure 13: Bland-Altman plot illustrating agreement for User A at 90° flexion with valgus force.

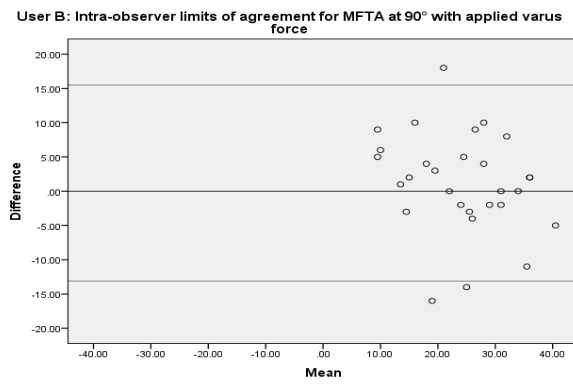


Figure 14: Bland-Altman plot illustrating agreement for User B at 90° flexion with varus force



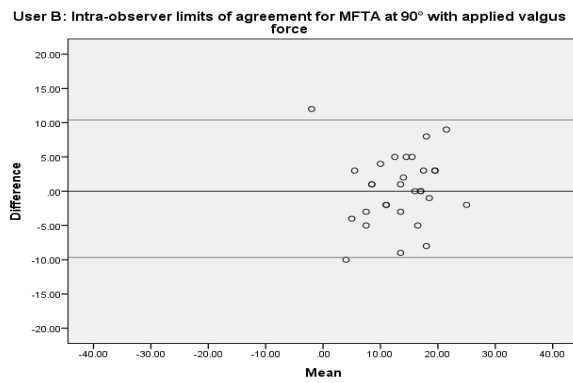


Figure 15: Bland-Altman plot illustrating agreement for User B at 90° flexion with valgus force.

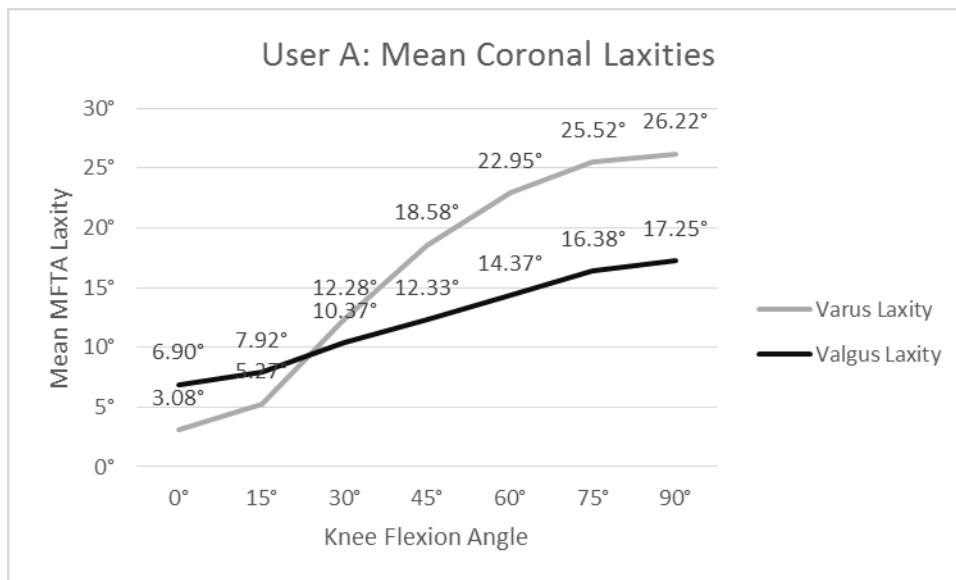


Figure 16: Mean angulation in the coronal plane with applied force observed for User A.

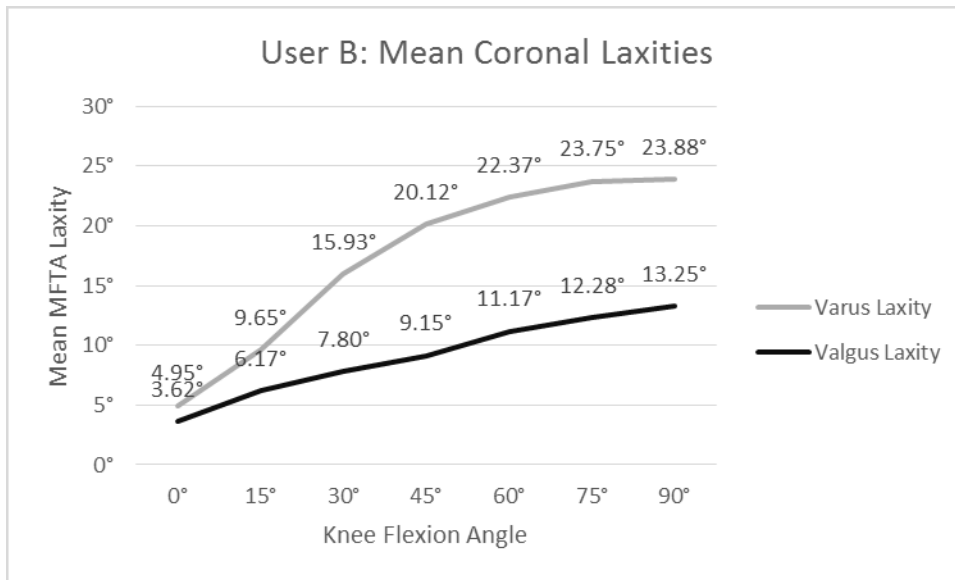


Figure 17: Mean angulation in the coronal plane with applied varus or valgus force observed for User B.

	<b>User A</b>	<b>User B</b>
<b>Supine MFTA at extension</b>	0.33° varus	0.13° valgus
<b>Bipedal MFTA</b>	0.43° varus	0.33° varus
<b>Monopedal MFTA</b>	0° (neutral)	0.5° varus

Table 1: Mean MFTA values calculated for each user.

	<b>User A CR</b>	<b>User B CR</b>	<b>ICC</b>
<b>Supine MFTA at Extension</b>	1.55	1.33	0.67
<b>Bipedal MFTA</b>	4.02	6.43	0.32
<b>Monopodal MFTA</b>	5.17	6.52	0.54

Table 2: Intra-observer CRs and inter-observer ICC calculated for supine MFTA at extension (initial registration MFTA with no force applied), bipedal MFTA in extension and monopodal MFTA in extension

<b>Flexion Angle</b>	<b>Applied Varus Force</b>	<b>Applied Valgus Force</b>
0°	3.02	4.42
15°	4.51	5.69
30°	5.84	8.13
45°	11.36	7.92
60°	11.90	7.39
75°	9.60	5.95
90°	8.00	8.10

Table 3: Intra-observer CR calculated for User A with applied varus and valgus force through the range of flexion.

<b>Flexion Angle</b>	<b>Applied Varus Force</b>	<b>Applied Valgus Force</b>
0°	5.08	4.74
15°	5.70	7.07
30°	6.79	6.97
45°	11.45	9.41
60°	14.19	5.80
75°	15.00	5.74
90°	15.78	6.23

Table 4: Intra-observer CR calculated for User B with applied varus and valgus force through the range of flexion.

<b>Flexion Angle</b>	<b>Applied Varus Force</b>	<b>Applied Valgus Force</b>
0°	0.57	0.54
15°	0.45	0.50
30°	0.61	0.72
45°	0.57	0.61
60°	0.64	0.61
75°	0.51	0.38
90°	0.65	0.61

Table 5: Inter-observer ICC for applied varus and valgus force through the range of flexion.



	<b>User A</b>	<b>User B</b>
Supine MFTA at Extension	0.126	0.1
Bipedal MFTA	0.203	0.495
Monopodal MFTA	0.161	0.815
0° varus	0.441	0.03
0° valgus	0.184	0.066
15° varus	0.418	0.187
15° valgus	0.383	0.12
30° varus	0.039	0.148
30° valgus	0.383	0.034
45° varus	0.022	0.085
45° valgus	0.09	0.046
60° varus	0.018	0.059
60° valgus	0.017	0.071
75° varus	0.012	0.034
75° valgus	0.008	0.097
90° varus	0.028	0.003
90° valgus	0.003	0.094

Table 6: Mann-Whitney U-test p-values for effect of sex on trial outcomes. This illustrates a tendency for a difference with sex observed with increasing flexion, however the results are not equivocal for both users.

	<b>User A</b>	<b>User B</b>
Supine MFTA at Extension	0.241	0.292
Bipedal MFTA	0.022	0.181
Monopodal MFTA	-0.008	0.098
0° varus	0.23	0.403 (p=0.027)
0° valgus	-0.329	-0.372 (p=0.043)
15° varus	-0.073	-0.001
15° valgus	-0.371 (p=0.043)	-0.283
30° varus	-0.232	-0.123
30° valgus	0.01	-0.321
45° varus	-0.226	-0.146
45° valgus	-0.16	-0.214
60° varus	-0.105	-0.087
60° valgus	-0.135	-0.034
75° varus	-0.252	-0.071
75° valgus	-0.243	-0.068
90° varus	-0.158	-0.101
90° valgus	-0.18	-0.105

Table 7: Spearman rank correlation coefficients for effect of BMI on measured outcomes. The largely negative values suggest a tendency for coronal laxity in flexion to decrease if BMI is increased, although this is not entirely uniform and is rarely at a level of statistical significance.

	<b>User A</b>	<b>User B</b>
Supine MFTA at Extension	0.254	0.088
Bipedal MFTA	0.119	-0.075
Monopedal MFTA	0.118	-0.083
0° varus	0.235	0.312
0° valgus	-0.408 (p=0.025)	-0.2
15° varus	-0.068	-0.157
15° valgus	-0.33	-0.169
30° varus	-0.045	-0.105
30° valgus	-0.028	-0.218
45° varus	-0.379 (p=0.039)	-0.15
45° valgus	-0.177	-0.201
60° varus	-0.202	-0.123
60° valgus	-0.091	0.001
75° varus	-0.299	-0.102
75° valgus	-0.095	0.023
90° varus	-0.149	-0.116
90° valgus	-0.015	0.01

Table 8: Spearman rank correlation coefficients for effect of age on measured outcomes. Largely negative values in flexion suggest a tendency for coronal laxity to decrease with increasing age, although this is not uniform.

