ORIGINAL ARTICLE

Virtual 3D femur model to assess femoral version: comparison to the 2D axial slice approach

Jade Iwasaka-Neder[1] · Sarah D Bixby[1] · M Alejandra Bedoya[1] · Enju Liu[2] · Delma Y Jarrett[1,3] · Donna Agahigian[1] · Andy Tsai[1]

Abstract

Background Quantifying femoral version is crucial in diagnosing femoral version abnormalities and for accurate pre-surgical planning. There are numerous methods for measuring femoral version, however, reliability studies for most of these methods excluded children with hip deformities.

Objective To propose a method of measuring femoral version based on a virtual 3D femur model, and systematically compare its reliability to the widely used Murphy's 2D axial slice technique.

Materials and methods We searched our imaging database to identify hip/femur CTs performed on children (<18 years old) with a clinical indication of femoral version measurement (September 2021—August 2022). Exclusion criteria were prior hip surgery, and inadequate image quality or field-of-view. Two blinded radiologists independently measured femoral version using the virtual 3D femur model and Murphy's 2D axial slice method. To assess intrareader variability, we randomly selected 20% of the study sample for re-measurements by the two radiologists >2 weeks later. We analyzed the reliability and correlation of these techniques via intraclass correlation coefficient (ICC), Bland-Altman analysis, and deformity subgroup analysis.

Results Our study sample consisted of 142 femurs from 71 patients (10.6±4.4 years, male=31). Intra- and inter-reader correlations for both techniques were excellent (ICC≥0.91). However, Bland-Altman analysis revealed that the standard deviation (SD) of the absolute difference between the two radiologists for the Murphy method (mean 13.7°) was larger than that of the 3D femur model technique (mean 4.8°), indicating higher reader variability. In femurs with hip flexion deformity, the SD of the absolute difference for the Murphy technique was 17°, compared to 6.5° for the 3D femur model technique. In femurs with apparent coxa valga deformity, the SD of the absolute difference for the Murphy technique was 10.4°, compared to 5.2° for the 3D femur model technique.

Conclusion The 3D femur model technique is more reliable than the Murphy's 2D axial slice technique in measuring femoral version, especially in children with hip flexion and apparent coxa valga deformities.

✉ Jade Iwasaka-Neder
jade.iwasaka-neder@childrens.harvard.edu

1 Department of Radiology, Boston Children's Hospital, 300 Longwood Ave, Boston, MA 02115, USA

2 Department of Pediatrics, Boston Children's Hospital, 300 Longwood Ave, Boston, MA 02115, USA

3 Department of Radiology, New York Presbyterian Hospital-Weill Cornell, 525 East 68th Street, New York, NY 10065, USA

Introduction

Femoral version (or torsion) is defined as the angle between the longitudinal axis of the femoral neck and the posterior margins of the femoral condyles, as measured on its 2D projection onto the plane perpendicular to the femoral shaft [1–3]. By convention, positive and negative version angles correspond to external and internal rotation of the femoral neck relative to the condyles, respectively. These rotations are commonly referred to as femoral anteversion and retroversion, respectively [4, 5]. Radiologists and orthopedists routinely assess femoral

version for the diagnosis of femoral torsional abnormalities and for pre-surgical planning [6–8].

In the context of assessing femoral version, the ground truth had previously been established in the literature utilizing femoral specimens [4, 9]. However, an important factor in the accuracy of femoral anteversion measurements is patient positioning, including rotation and hip and knee flexion deformities, which are commonly observed in patients with cerebral palsy [10, 11]. Furthermore, femoral version measurement accuracy is affected by the presence of femoral head and neck deformities, as often seen in conditions such as avascular necrosis of the femoral head and coxa valga [11, 12]. To realistically assess the effects of positioning and deformities on version measurements, methods should be compared through studies conducted on living children. This precludes obtaining ground truth measurements via cadaveric studies, and comparison should instead be made to the clinical reference standard.

Various techniques based on computed tomography (CT), magnetic resonance imaging (MRI), ultrasound, fluoroscopy, and biplane radiography have been proposed to quantify femoral version [2, 13–15]. Each technique specifies certain anatomical landmarks to guide how the femoral version is determined [2, 13–15]. Unfortunately, morphologic abnormalities of the femur or alterations in patient positioning may create ambiguity in identifying these landmarks, affecting our ability to assess femoral version [16–18].

Popular MRI- and CT-based techniques for measuring femoral version mostly utilize 2D axial slices for measurements, with reports of high intra- and inter-observer agreements [13, 19, 20]. Among the methods available for measuring femoral version, the 2D axial slice method proposed by Murphy et al. stands out, as it is the first CT-based method to be validated through accuracy analysis using cadaveric femurs [1]. To date, the Murphy method is one of the most widely utilized approaches for assessing femoral version, and is considered by many as the clinical reference standard [21–26]. However, the study by Murphy et al., along with many other femoral version reliability studies, excluded femurs with deformities such as coxa valga and fixed hip flexion deformity [27–29], or did not evaluate how these deformities influence the femoral version measurements, making their results less applicable to the same population of patients who are most at risk of having abnormal femoral version.

In this study, we describe a novel method of measuring femoral version based on a virtual 3D model of the femur. We compare this novel method to the widely-used 2D axial slice method proposed by Murphy et al. [1], both in terms of reader reliability and generalizability.

Materials and methods

This retrospective study was approved by our institutional review board. It was compliant with the Health Insurance Portability and Accountability Act, and informed consent was waived.

Study population

We conducted a computerized search of the image archive at our large tertiary children's hospital to identify hip/femur CTs performed on children (<18 years old) for the clinical indication of femoral version measurements (September 2021—August 2022). Additional inclusion criteria were: 1) utilization of the low-dose CT image acquisition protocol as described below, and 2) image acquisition on one of two Dual-Source CT scanners using Turbo Flash mode (Siemens Somatom Force, Erlangen, Germany). These are the scanners most commonly used at our institution for these examinations. Exclusion criteria were: 1) prior hip or femur surgery, and 2) inadequate image quality (i.e. blurry images secondary to motion) or inadequate imaging field-of-view (FOV) that did not include the entire femur.

CT imaging protocol

All CTs were performed on dual-energy Siemens Somatom Force scanners. The imaging protocol was optimized to achieve the lowest radiation exposure while maintaining sufficient image quality to enable generation of high-quality 3D femur models for visualization and femoral version measurements. We used weight-adjusted kVp (with reference kVp of 100 and 120 for patients <55 kg and ≥55 kg, respectively); and automated tube current modulation (with quality reference mAs of 150 and 40 for patients <55 kg and ≥55 kg, respectively). Dual source scanning was utilized with helical pitch of 3. The imaging FOV was from the top of the iliac crests through the knees in order to capture the entire femur for 3D modeling. The volume CT dose index ($CTDI_{vol}$) and the dose-length product (DLP) were recorded as estimates of the radiation dose for each CT study.

Virtual 3D femur model

We used a 3D workstation with multi-planar reconstruction capabilities (Synapse, Fujifilm, Japan) to generate virtual 3D femur models. Each model was generated semi-automatically from the axial CT images via a combination of intensity thresholding and edge-preserving smoothing, followed by manual fine-tuning to ensure biofidelity. The proximal third of the model (including the femoral head and neck)

was assigned an opaque color, while the distal two thirds (including the femoral condyles) were assigned a transparent color (Fig. 1). When the proximal and the distal extremities of the 3D femur model overlap during spatial rotation of the model, this variable shading allows the users to see through the femoral condyles and evaluate the femoral head and neck. The virtual 3D model was then oriented vertically with the posterior margins of the femoral condyles aligned to the coronal plane. From there, we generated a tumble sequence of the 3D model (in 2.5° increments) by spatially rotating it 360° in the sagittal plane centered about the mid femoral shaft. This virtual 3D femur model, including its 360° tumble, was imported to the hospital picture archiving and communication system (PACS).

To measure the femoral version using PACS, the virtual 3D femur model was rotated in the sagittal plane until the femoral shaft was viewed along its long axis (i.e., looking straight down the femoral shaft) (Fig. 1, Online Resource 1). In this position, the femoral condyles and the femoral neck are superimposed on top of one another. Depending on whether the model was rotated 90° or 270° in the tumble sequence, the axial view of the femur would be either from a caudal-to-cranial or from a cranial-to-caudal point-of-view. A single angle was then drawn between the femoral neck axis and the posterior margins of the femoral condyles, and we denoted this as the angle of femoral version.

Murphy's 2D axial slice method

The Murphy technique of measuring femoral version relied strictly on 2D axial CT images to measure two separate angles—the proximal and distal angles—which were combined to determine femoral version (Fig. 2, Online Resource 2) [1]. Specifically, for the proximal angle, a line was drawn from the center of the femoral head on one axial CT image to the femoral neck base at the level of the lesser trochanter in a different axial CT image. The proximal angle was formed by this line and the horizontal plane. For the distal angle, the axial CT image containing the posterior margins of the femoral condyles was identified, and the distal angle was taken between this posterior femoral condylar line and the horizontal plane. If the femoral condyles were externally rotated, the distal angle was subtracted from the proximal angle to yield the femoral version angle. Alternatively, if the femoral condyles were internally rotated, the distal angle was added to the proximal angle to yield the femoral version angle.

Fig. 1 Example showing how femoral version was measured via the 3D femur model methodology in a 16-year-old boy. From left to right, the entire model was rotated in the sagittal direction to allow an axial perspective of both femoral condyles and the femoral neck. The user has the option of viewing the femur from a cranial-to-caudal (**a**), or from a caudal-to-cranial (**b**) point-of-view. The center of rotation was the mid femoral shaft. A single angle was drawn between the femoral neck axis and the posterior aspect of the femoral condyles. In this particular example, Reader #1 and #2 measured the femoral version as 7° and 8°, respectively
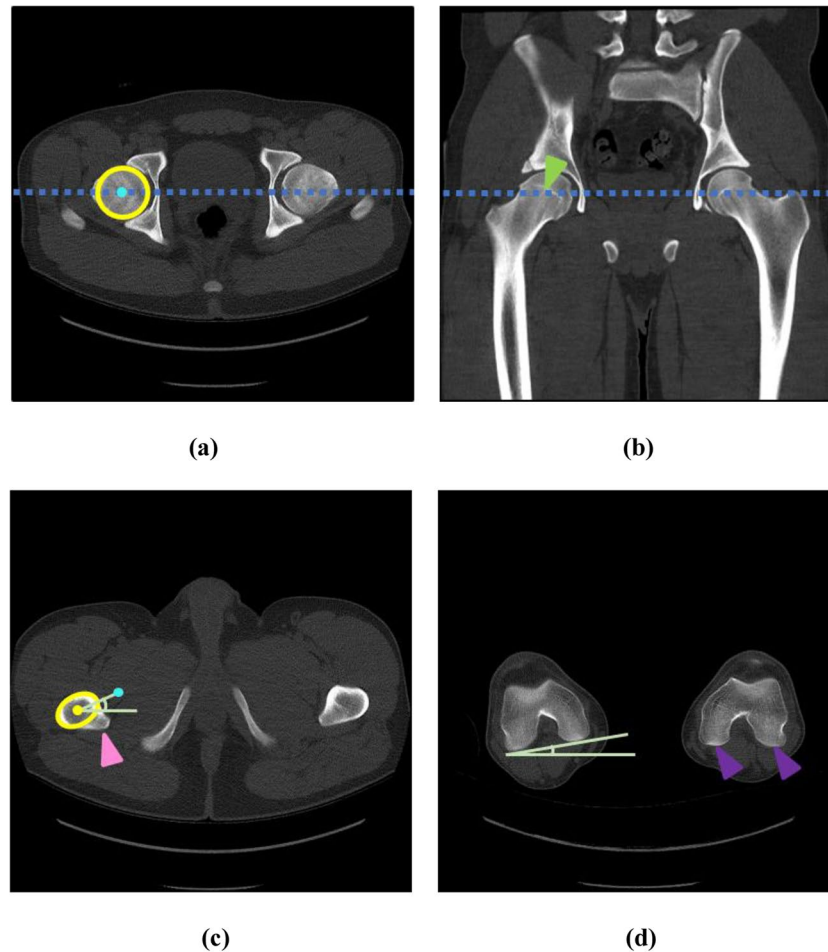

(a)

(b)

(a)

(b)

(c)

(d)
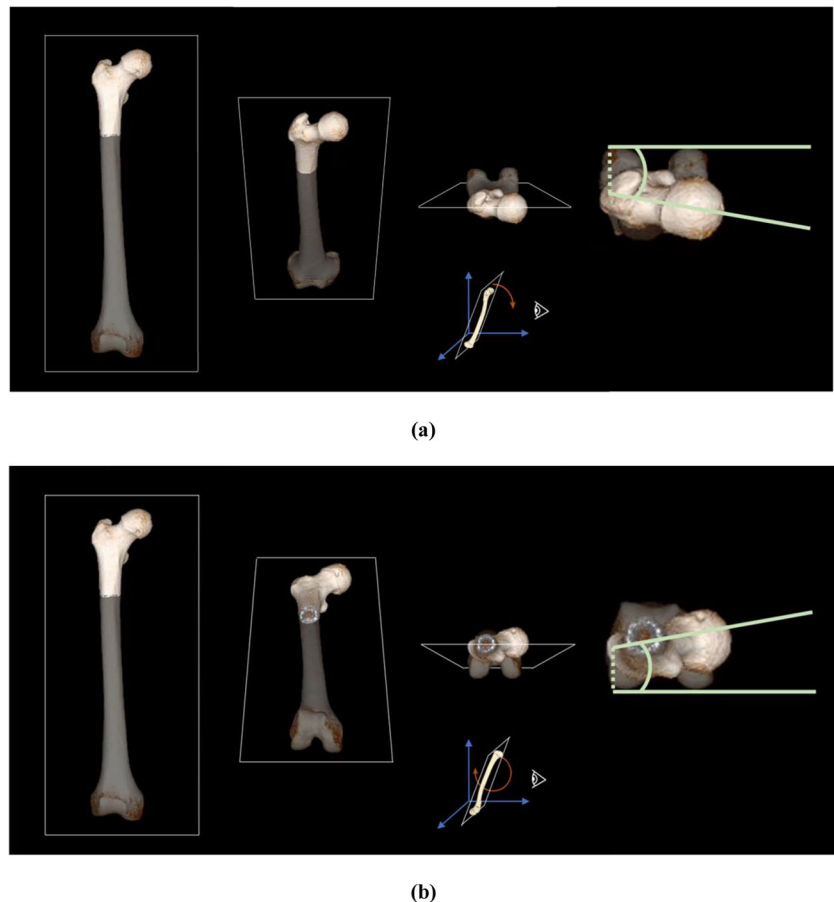
Fig. 2 Femoral version measurement via Murphy's 2D axial slice methodology in a 16-year-old boy (same patient as in Fig. 1). The center of the femoral head on the axial plane (blue dot in **a**) is determined by triangulation with the center of the femoral head on the coronal images (green arrowhead in **b**). The cursor is placed at the center of femoral head on the axial plane (blue dot in **a**) while the CT images are advanced caudally to the base of the femoral neck (yellow dot in **c**), which is defined as at the level of the lesser trochanter (pink arrowhead in **c**). The proximal angle is formed by the line connecting these two points and the horizontal plane (angle in **c**). The axial CT images demonstrating the posterior aspects of the femoral condyles are identified (purple arrowheads in **d**). The distal angle is drawn between a line tangential to the posterior margin of the femoral condyles and the horizontal plane (angle in d). In this particular example, there was external rotation of the femoral condyles, and therefore, the distal angle was subtracted from the proximal angle to determine the overall angle of torsion. In this example, Reader #1 and #2 measured the femoral version as 7° and 8°, respectively

Image interpretation

In a blinded randomized fashion, two board-certified pediatric radiologists (S.D.B. with 14 years, and M.A.B. with 1 year of post-fellowship experience), both with subspecialty expertise in musculoskeletal imaging, independently measured the version angles of the right and left femurs within the study sample, using electronic measuring tools available in PACS. Two anonymized worklists were assembled in PACS, one containing only axial CT slices and the other only 3D femur models. Each radiologist measured the version angle using the two different techniques: Murphy's 2D axial slice and the virtual 3D femur model method. Only one technique was utilized in each reading session so as to minimize bias.

The sessions for each technique were conducted within the same week, and the measurements documented by a research fellow (J.I.N). To assess intra-reader variability, 20% of the CT exams were randomly selected and their femoral versions were re-measured by the two radiologists (using both measurement techniques) >2 weeks later. The re-measurements were also conducted on the anonymized PACS worklists in a blinded randomized independent fashion. Each reader only utilized one technique in each re-reading session to avoid measurement bias. The sessions for the re-measurements were conducted within the same week, and these measurements were again documented by a research fellow (J.I.N).

To determine femoral coxa valga and hip flexion deformity, a research fellow (J.I.N.) independently measured the

apparent neck-shaft angle and the degree of hip flexion for each hip/femur, using reformatted 2D coronal and sagittal CT images, respectively (Figs. 3 and 4). The measurements were conducted using electronic measuring tools available in PACS. Hip flexion deformity was defined as having a hip flexion angle of >10° [30, 31]. Coxa valga deformity was defined as having an apparent femoral neck-shaft angle of >140° [32–34]. We used apparent neck-shaft angle in our study because the true neck-shaft angle is the degree of femoral neck inclination measured on a AP radiograph or true coronal CT scan when the femoral neck is positioned parallel to the horizontal plane. Since we did not correct for the femoral neck positioning in each patient, the femoral neck-shaft angle that we measured is customarily referred to as the "apparent" femoral neck-shaft angle [35]. We chose to classify coxa valga deformity based on *apparent* (rather than true) femoral neck-shaft angle for the following reasons: 1) we cannot control leg positioning in our patients with hip deformities; 2) the published reference values of femoral neck-shaft angle were drawn from studies that also used apparent neck-shaft angles [36–38]; and 3) true

femoral-neck angle calculation would entail correcting for leg positioning and femoral version deformities [35, 39–42], and thus would undermine our study (resulting in circular reasoning).

Statistical analyses

We first analyzed and compared the reliability of these two methods (virtual 3D femur model and Murphy's 2D axial slice technique) via intraclass correlation coefficient (ICC), Bland-Altman analysis, and scatterplots, with additional subgroup analysis to assess the effects of hip and femur deformities [43–45]. For each femur, the mean and standard deviation (SD) of the absolute difference between the average reader measurements from the two methods were calculated. We used mixed effects model to examine the relationship between the two methods accounting for the correlation between the left and right femurs of the same patient. Apparent coxa valga and hip flexion deformity were included as covariates in the model. We used the following guidelines for the interpretation of ICC: <0.40, 0.40-0.59,
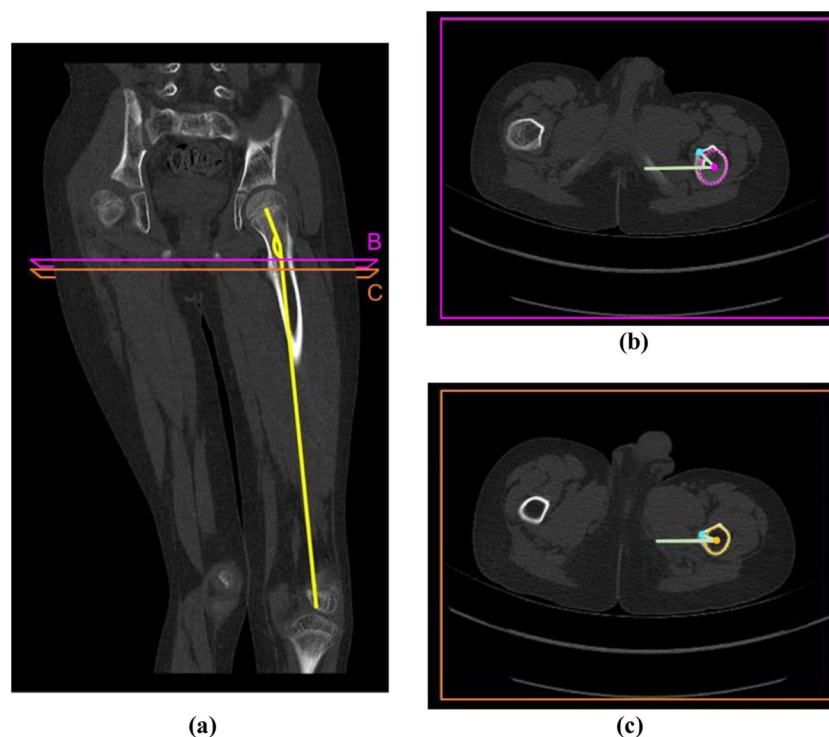

Fig. 3 Coronal (**a**) and axial images (**b** and **c**) of a 10-year-old boy with 153° of apparent neck-shaft angle on the left leg (yellow angle in **a**) and 154° on the right leg (not shown). The apparent neck-shaft angle was measured between the longitudinal axis of the femoral neck and the femoral shaft. The two axial images outlined in pink (**b**) and orange (**c**) correspond to different axial slices where the lesser trochanter would be considered as most prominent by triangulation with the coronal image (B and C labels on **a**), making them both appropriate slices to use in measuring the femoral neck angle.

The blue dot represents the center of the femoral head in both images. Note the differing measurements of the proximal angles (olive color) when assessed based on these two axial slices. This is a problem in any 2D axial slice method of measuring femoral version, as the anatomy of the proximal femur is distorted in patients with apparent coxa valga, resulting in ambiguity of the lesser trochanter. In this example, Reader #1 and #2 measured the femoral version as 12° and 24°, respectively
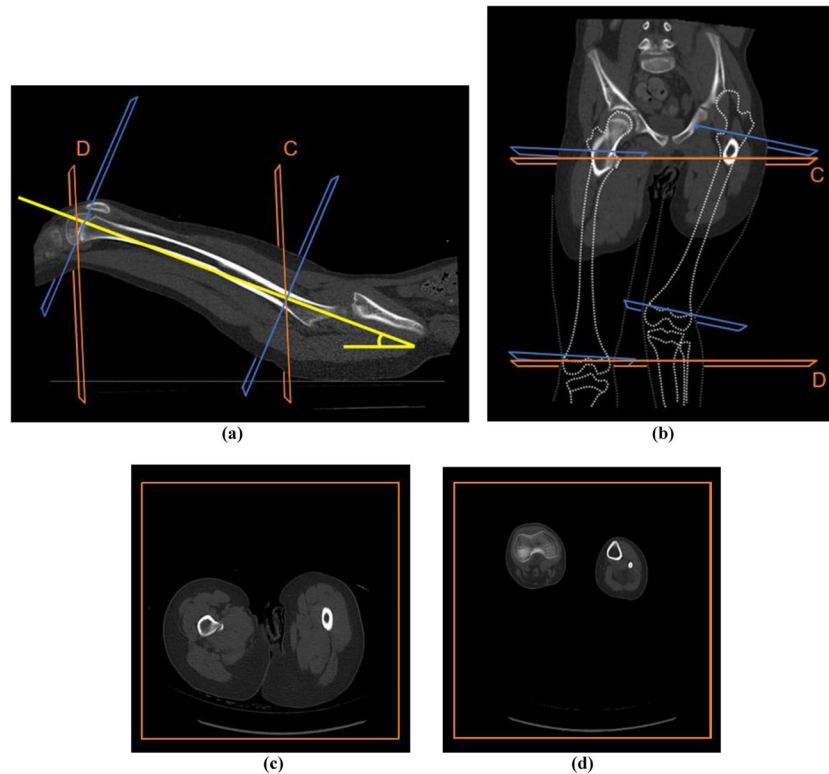
Fig. 4 Sagittal (**a**), coronal (**b**) and axial images (**c** and **d**) of a 9-year-old girl with 22° of hip flexion deformity on the right leg (yellow angle on a) and 39° on the left leg (not shown). The flexion angle was measured between the longitudinal axis of the femoral shaft and the horizontal plane (as represented by the outline of the CT table on **a**). On the coronal image (**b**), dashed lines illustrate how the legs and bones were positioned at the time of the CT acquisition. The 2D planes outlined in orange (C and D labels on **a** and **b**) correspond to the axial images (**c** and **d**) the radiologists used to assess the femoral version via the 2D axial slice technique. The 2D planes outlined in blue represent what would be the axial images if the scanning plane was precisely along the short axis of the femoral shaft (i.e., if the patient was lying flat on the scanner). The hip flexion deformity caused the femur to hinge anteriorly at the level of the hip, reorienting the femur into an oblique lie (and no longer flat to the CT table). Hence, when the radiologist scrolls through the axial CT images, each axial slice is no longer axial to the femoral shaft. Selecting the most appropriate slice according to the presence of landmarks becomes more challenging and subjective, and affects the femoral version measurements and negatively influences the agreement. In this example, Reader #1 and #2 measured the femoral version on the left leg as -57° and -33°, respectively

0.60-0.74, and 0.75-1.0 as poor, fair, good, and excellent correlations, respectively [46]. All statistical analyses were conducted with a statistical significance of $P<0.001$, using MATLAB 8.3 (MathWorks, Natick, MA) and *SAS* 9.4 (*SAS* Institute, Cary, NC).

Results

Study sample characteristics

Our study sample consisted of 71 CTs from 71 different patients (male=31, female=40) (Fig. 5). The clinical diagnosis that prompted femoral version assessment in this sample encompassed a diverse group of patients. It included 26 patients with cerebral palsy, 26 patients with developmental dysplasia of the hip (either idiopathic or secondary to various syndromes such as Down syndrome, Noonan syndrome,
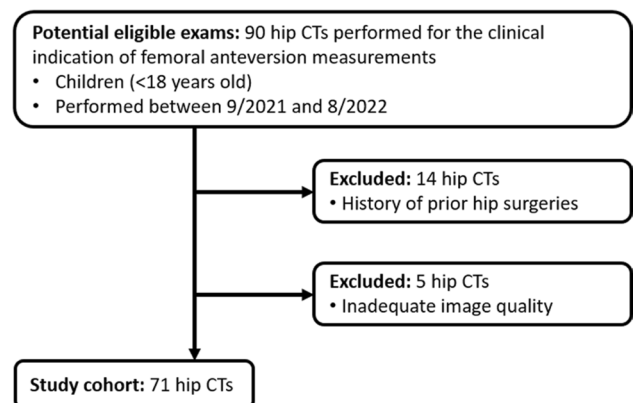

Fig. 5 Study flowchart illustrating the identification of the study sample

congenital cytomegalovirus infection, genetic leukodystrophy, among others), 11 patients with femoroacetabular impingement, 4 patients with Perthes disease, 3 patients with spinal muscular atrophy, and 1 patient with chronic hip pain. The average age at the time of the clinical CT examination was 10.6 ± 4.4 years (range=4.0-17.9 years). For the 71 CTs, the mean and SD of the $CTDI_{vol}$ were 1.46 mGy and 0.9 mGy, respectively, with range of 0.33 to 4.35 mGy; and the mean and SD of the DLP were 84.16 mGy·cm and 53.57 mGy·cm, respectively, with range of 10.69 to 238.09 mGy·cm. The 71 CTs yielded 142 femurs (two femurs per CT). Of the 142 femurs, 81 (57%) had at least one of the investigated deformities. Table 1 provides additional demographic details regarding our study sample.

Reader agreement

The intra-reader correlations for the two radiologists when using the 2D axial slice method were excellent (ICC=0.96, with 95% confidence interval (CI): [0.92, 0.98]; and 0.99, with 95% CI: [0.97, 0.99]). Similarly, the intra-reader correlations for the two radiologists when using the 3D femur model method were excellent (ICC=0.91, with 95% CI: [0.86, 0.95]; and 0.97, with 95% CI: [0.93, 0.98]).

When the entire study sample was analyzed (N=142), the inter-reader correlation for both methods was excellent (ICC≥0.95, Table 2, Fig. 6). For the 2D axial slice method, Bland-Altman analysis showed that the mean and SD of the absolute difference of measurements by the two radiologists were 1.2° and 13.7°, respectively; while those for the 3D femur model method were -1° and 4.8°, respectively (Fig. 7).

When femurs without hip flexion or apparent coxa valga deformities were analyzed (N=61), the inter-reader correlation for both methods was excellent (ICC≥0.94, Table 2, Fig. 6). Using the 2D axial slice method, Bland-Altman analysis showed that the mean and SD of the absolute difference of measurements by the two radiologists were -0.8° and 3.6°, respectively; while those for the 3D femur model method were -0.7° and 4.4°, respectively (Fig. 7).

When femurs with only apparent coxa valga deformities were analyzed (N=41), the inter-reader correlation for both methods was excellent (ICC≥0.94, Table 2, Fig. 6). Using the 2D axial slice method, Bland-Altman analysis showed that the mean and SD of the absolute difference of measurements by the two radiologists were -0.5° and 10.4°, respectively; while those for the 3D femur model method were 1.5° and 5.2°, respectively (Fig. 7).

When femurs with only hip flexion deformities were analyzed (N=6), the inter-reader correlation for both methods was excellent (ICC≥0.78, Table 2, Fig. 6). Using the 2D axial slice method, Bland-Altman analysis showed that the mean and SD of the absolute difference of measurements by the two radiologists were -0.3° and 16.9°, respectively; while those for the 3D femur model method were 1.3° and 6.5°, respectively (Fig. 7).

When femurs with both hip flexion and apparent coxa valga deformities were analyzed (N=34), the inter-reader correlation for both methods was excellent (ICC≥0.89, Table 2, Fig. 6). Using the 2D axial slice method, Bland-Altman analysis showed that the mean and SD of the absolute difference of measurements by the two radiologists were 5.5° and 25.4°, respectively; while those for the 3D femur model method were -1.1° and 5.3°, respectively (Fig. 7).

Comparison of the two different methods

Multivariable mixed effect models showed that the relationship between the difference and the mean femoral version measurements of the two methods was dependent on hip and femur deformity status. Specifically, the slopes of

Table 1 Descriptive characteristics of the study sample by subgroups of hip deformities. Unless otherwise indicated, data are medians; data in parentheses are the interquartile range

	All femurs	Femurs without hip flexion or apparent coxa valga deformity	Femurs with only apparent coxa valga deformity	Femurs with only hip flexion deformity	Femurs with hip flexion and apparent coxa valga deformities
No. of femurs	142	61	41	6	34
Sex					
M ‡	62 (44)	21 (34)	19 (46)	2 (23)	20 (59)
F ‡	80 (56)	40 (66)	22 (54)	4 (67)	14 (41)
Age (y) †	10.6 ± 4.4	13.2 ± 4.2	9.2 ± 3.7	10 ± 2	7.5 ± 2.7
Femoral inclination angle (degrees)	141.5 (135 to 152)	134 (131 to 136)	150 (144.7 to 156)	137.5 (135 to 139)	152 (147 to 161)
Hip flexion angle (degrees)	2 (-1 to 14)	0 (-2 to 2)	3 (-1.2 to 6.2)	20 (12 to 31)	23 (19 to 32)

† Data are mean ± standard deviation

‡ Data are number of patients, and data in parentheses are percentages

Table 2 Inter-reader agreement between two readers measuring femoral versions via the two methods: 2D axial slice method and the virtual 3D femur model method

Technique		All femurs	Femurs without hip flexion or apparent coxa valga deformity	Femurs with only apparent coxa valga deformity	Femurs with only hip flexion deformity	Femurs with both hip flexion and apparent coxa valga deformity
	No. of femurs	142	61	41	6	34
2D Axial Slice	Inter-reader ICC	0.95	0.97	0.94	0.78	0.89
	[95% CI]	[0.93, 0.96]	[0.94, 0.98]	[0.88, 0.96]	[-0.01, 0.97]	[0.79, 0.94]
	P-value	*P*<0.001	*P*<0.001	*P*<0.001	*P*=0.03	*P*<0.001
	Bland-Altman analysis					
	Absolute bias (mean±SD)	1.17 ± 13.68	-0.75 ± 3.62	-0.54 ± 10.39	-0.33 ± 16.96	5.53 ± 25.37
	Limits of agreement	-26.2 to 28.53	-7.1 to 6.49	-21.32 to 20.25	-34.25 to 33.58	-42.46 to 53.52
Virtual 3D Femur Model	Inter-reader ICC	0.96	0.94	0.95	0.89	0.96
	[95% CI]	[0.94, 0.97]	[0.90, 0.96]	[0.90, 0.97]	[0.46, 0.98]	[0.91, 0.98]
	P-value	*P*<0.001	*P*<0.001	*P*<0.001	*P*=0.01	*P*<0.001
	Bland-Altman analysis					
	Absolute bias (mean±SD)	-1.03 ± 4.85	-0.69 ± 4.44	1.46 ± 5.15	1.33 ± 6.47	-1.09 ± 5.28
	Limits of agreement	-10.73 to 8.66	-9.57 to 8.2	-8.84 to 11.77	-8.94 to 11.60	-11.65 to 9.48

ICC intraclass correlation coefficient, *CI* confidence interval, *SD* standard deviation

the regression lines were -0.39, -0.75, -1.08, and -1.47 for femurs without hip flexion or apparent coxa valga, only with apparent coxa valga, only with hip flexion, and with both apparent coxa valga and hip flexion, respectively ($P<0.001$). Among femurs without hip flexion or apparent coxa valga, the difference in measurements from the two methods was consistently small (near zero) across the range of version angle measurements. However, among femurs with various deformities, the difference was large, especially in those with both hip flexion and apparent coxa valga deformities, and at the negative values of mean version angle (i.e. retroversion) (Fig. 8).

When analyzing the entire study sample ($N=142$), the measurements made via the 3D femur model method and those made via the 2D axial slice method were not significantly correlated ($r=-0.02$, with 95% CI: [-0.18, 0.15], $P=0.86$). The mean and SD of the absolute difference of the two methods were 21.4° and 48°, respectively; and the limits of agreement were -72.7° to 115.5° (Fig. 8). However, when analyzing only femurs that did not have hip flexion or apparent coxa valga deformities ($N=61$), the measurements by the two techniques were highly correlated ($r=0.89$, with 95% CI: [0.82, 0.93], $P<0.001$). The mean and SD of the absolute difference by the two methods were -6.1° and 6.7°, respectively; and the limits of agreement were -19.3° to 7.0°. Based on subgroup analysis, deformities (apparent coxa valga and hip flexion deformities) are associated with decreased correlation between the two methods (Table 3). The same effects were observed for measurements of each of the two radiologists individually (Online Resources 4-6).

Discussion

Many of the 2D axial CT methods of measuring femoral version are associated with inherent limitations related to variabilities in how the measurements are acquired [16]. The two most obvious and consequential limitations include: (1) uncertainty in determining the appropriate 2D slices on the axial CT images to draw the femoral neck line; and (2) the need to prescribe the axial scanning plane precisely along the short axis of the femoral shaft [28]. To remedy these issues, we proposed a virtual 3D femur model method of measuring femoral version. This approach allows the radiologist to view the 3D morphology and alignment of the entire femur rather than inferring its complex shape from multiple 2D axial CT images, thus minimizing measurement errors and reader variability. Our reliability study showed that the inter-reader agreement for the virtual 3D femur model method was superior to the Murphy's 2D axial slice method. The excellent inter-reader agreement of our 3D femur model method was robust to apparent coxa valga and hip flexion deformities. In contrast, the inter-reader agreement of the 2D axial slice method decreased substantially when evaluating children with these deformities.

Both intra- and inter-reader ICCs were good to excellent when evaluating each of the two techniques of measuring femoral version that we investigated in this study, and are in keeping with the literature [13, 19, 20]. This could potentially create a false impression that the differences between measurements are consistently small [44, 47]. In fact, high ICC values simply reflect strong agreement and minimal
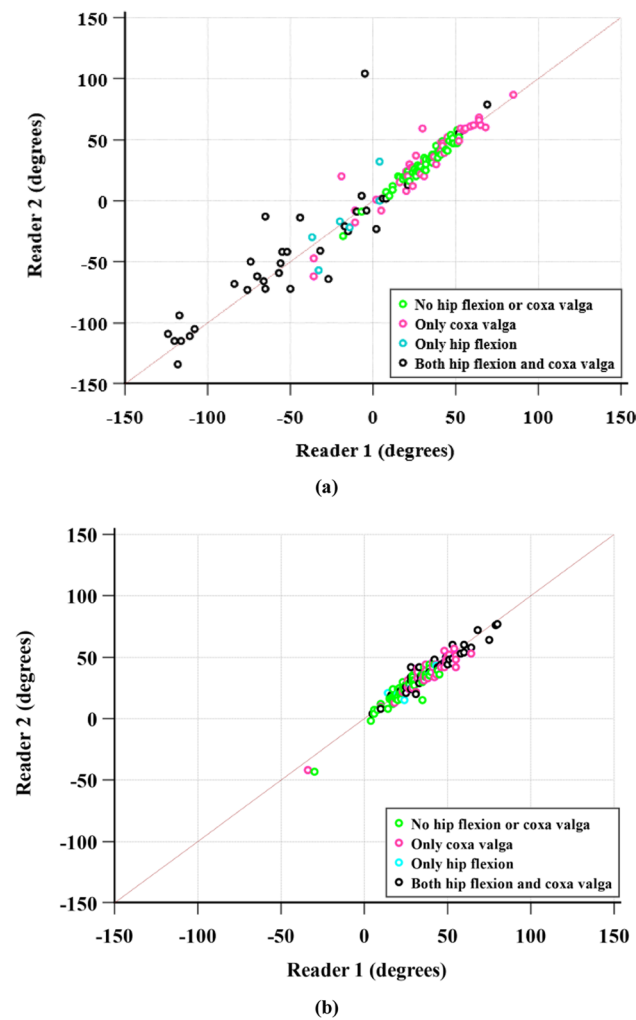
Fig. 6 Scatterplots comparing femoral version measurements by the two readers using Murphy's 2D axial slice method (**a**), and the virtual 3D femur model method (**b**). The brown line denotes the identity line (where the measurements by the two readers are identical). Each plot contains 142 data points. Green points correspond to femoral version measurements of femurs without hip flexion or apparent coxa valga deformities. Magenta points correspond to femoral version measurements of femurs with apparent coxa valga deformities. Cyan points correspond to femoral version measurements of femurs with hip flexion deformities. Black points correspond to femoral version measurements of femurs with both hip flexion and apparent coxa valga deformities. Note the greater dispersion of the data points from the identity line with deformities when using the 2D axial slice method (**a**)

Fig. 7 Bland-Altman plots comparing femoral version measurements generated by the two readers using Murphy's 2D axial slice method (**a**), and the virtual 3D femur model method (**b**). The axial slice method was associated with a mean bias of 1.2° and a standard deviation of 13.7° between the readers. The 3D femur model method was associated with a mean bias of -1.0° and a standard deviation of 4.8°. For the axial slice method, the limits of agreement (-26.2° to 28.5°) were wide, and those femurs with deformities were major contributors to this widening. In contrast, for the 3D femur model method, the limits of agreement (-10.7° to 8.7°) were much narrower, and those femurs with deformities had much less impact on the limits of agreement. Each plot contains 142 data points. Green points correspond to femoral version measurements of femurs without hip flexion or apparent coxa valga deformities. Magenta points correspond to femoral version measurements of femurs with apparent coxa valga deformities. Cyan points correspond to femoral version measurements of femurs with hip flexion deformities. Black points correspond to femoral version measurements of femurs with both hip flexion and apparent coxa valga deformities

variability between measurements made by two readers or two techniques, without necessarily considering the magnitude of the numerical differences between the measurements [45, 47]. Additional characterization of agreement through Bland-Altman difference analysis allows visualization of the magnitude of differences through the limits of agreement and SD [47], as illustrated by our results.

Many prior studies that investigated the reliability of femoral version measurements excluded femurs with hip
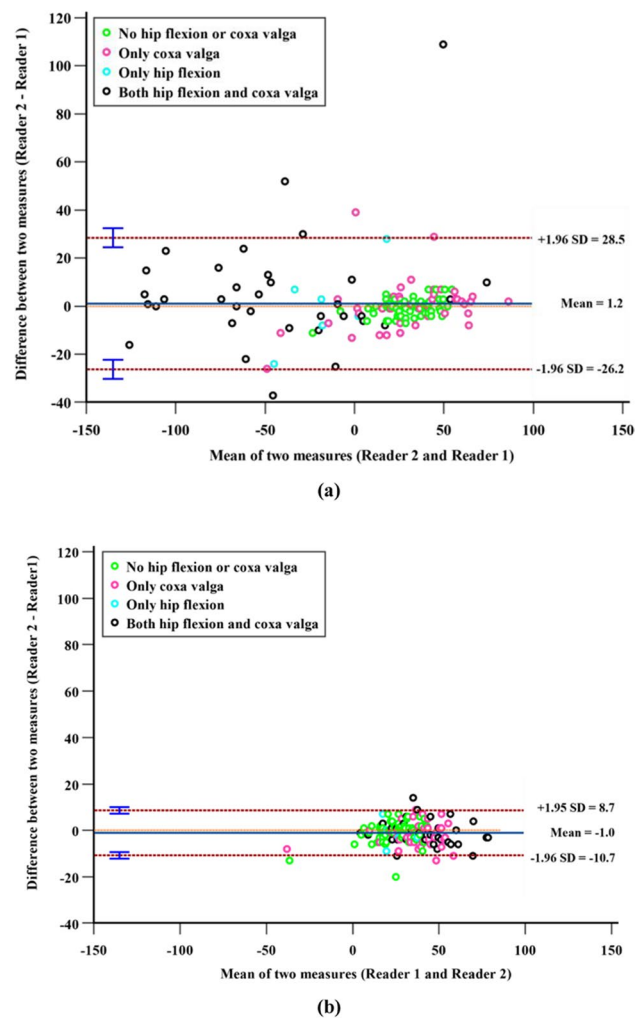
and femur deformities [27–29]. Ironically, one of the most common clinical indications for assessment of femoral version is the pre-operative planning of children with such deformities, as demonstrated by 57% of our sample

Fig. 8 Multivariable linear regression models (**a**), scatterplot (**b**), ▶ and Bland-Altman plot (**c**) showing the relationship between the two methods of measuring femoral version, as a function of deformities. On all plots, each data point corresponds to the average of the measurements made by Readers #1 and #2 for a given method. The orange line denotes the identity line (i.e., zero bias line, where the measurements by the two methods are identical). Green points and line (slope=-0.39) correspond to femoral version measurements of femurs without hip flexion or apparent coxa valga deformities. Magenta points and line (slope=-0.75) correspond to femoral version measurements of femurs with apparent coxa valga deformities. Cyan points and line (slope=-1.08) correspond to femoral version measurements of femurs with hip flexion deformities. Black points and line (slope=-1.47) correspond to femoral version measurements of femurs with both hip flexion and apparent coxa valga deformities (*P* for interaction <0.001). There is a decrease in correlation between the two methods in the subgroups with deformities (**b**). The mean difference between the two methods was 21.4°, with a standard deviation of 48°. The limits of agreement (-72.7° to 115.5°) between the two methods were wide, and those femurs with deformities contributed most to this widening (**c**)

(a)

(b)

(c)

presenting at least one of the examined deformities. Excluding them from reliability studies substantially diminishes the generalizability of these other techniques. In our investigation, we included patients with challenging deformities in our study sample to evaluate the generalizability of our methodology.

The median values of femoral version found with the 3D femur model method and the 2D axial slice method were similar in magnitude for femurs without deformity; however, they became increasingly discrepant in the presence of deformities. For instance, the median femoral version of subjects with both hip flexion and apparent coxa valgus deformity exhibited a difference of almost 90 degrees between the two methods (44 degrees according to the 3D femur model method versus -48 degrees according to the 2D axial slice approach). We hypothesize that this perceived discrepancy was partially due to projectional differences.

An increase in the hip flexion angle causes the femoral head to shift in relation to the base of the femoral neck (specifically, at the level of the lesser trochanter) (see Online Resource 3). When the long axis of the femoral shaft and the posterior femoral condylar line are parallel to the horizontal plane, an ideal axial CT image is obtained, allowing for accurate measurement of femoral version. Since femoral anteversion is the most common configuration of the femur, the femoral version is typically a positive angle, hence the positive values in the subgroups without flexion. However, when the hip is flexed, the lesser trochanter moves anteriorly relative to the femoral head, giving a false impression of femoral retroversion when assessed on 2D axial CT scans. This explains the negative values observed in the flexed subgroup of our analysis. Since the entire length of the femur is not visible on the 2D axial slices, it becomes challenging to perceive and account for this false impression
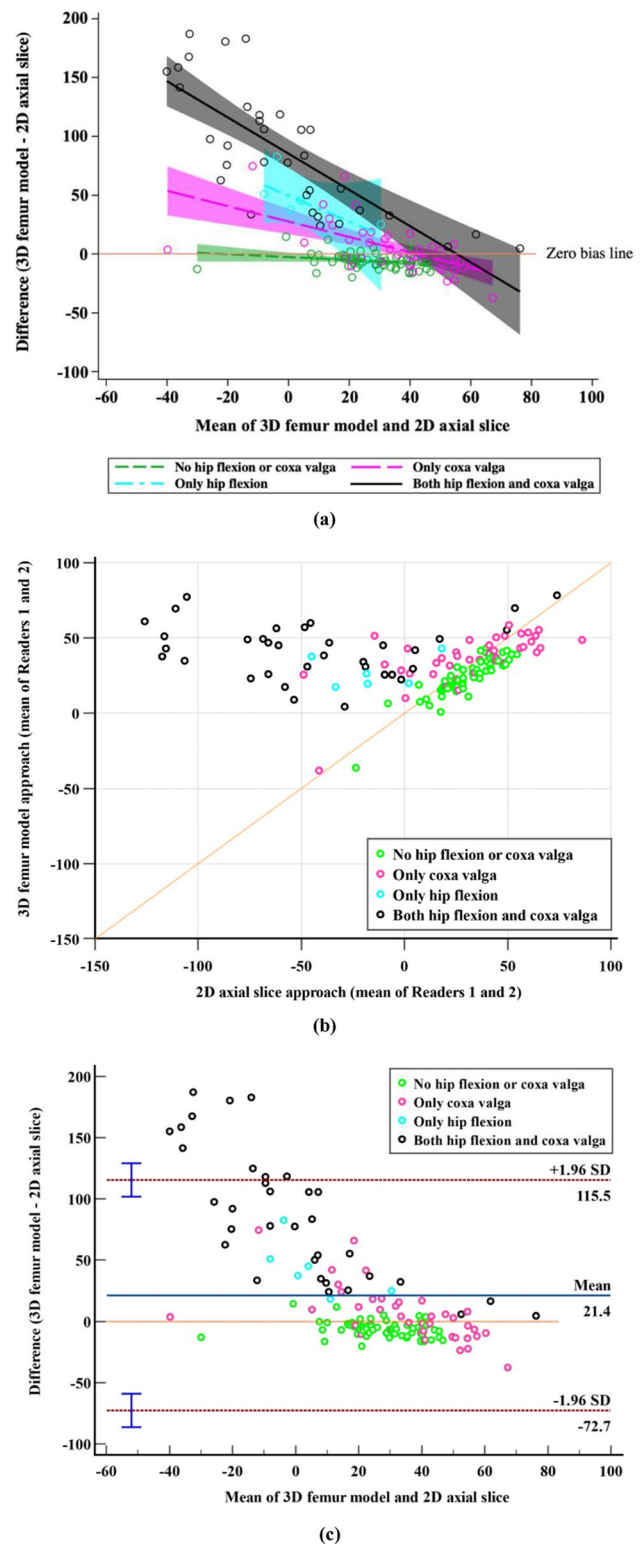
of retroversion. This pitfall in the 2D axial slice methods and the high discrepancy between the findings of the two methods in hip flexion deformity further supports the utility of our virtual 3D femur model in the assessment of femoral version in patients with hip deformities.

Table 3 Comparison between femoral version measurements obtained via the 2D axial slice method to those obtained via the virtual 3D femur model method, based on the average values of Readers #1 and #2. Results are presented by patient subgroup to characterize the effect of deformities on measurement agreement

	All femurs	Femurs without hip flexion or apparent coxa valga deformity	Femurs with only apparent coxa valga deformity	Femurs with only hip flexion deformity	Femurs with both hip flexion and apparent coxa valga deformity
No. of femurs	142	61	41	6	34
Femoral version based on 3D femur model method					
Median	32.2	24.5	38	23.2	44
Interquartile range	(21.5 to 43)	(17.5 to 32.5)	(30.7 to 47.7)	(19.5 to 37.5)	(29.5 to 55)
Femoral version based on 2D axial slice method					
Median	25	31	34	-18.25	-47.7
Interquartile range	(-8 to 41)	(22.1 to 40.4)	(17.4 to 51.9)	(-33.5 to 2)	(-74.5 to -9.5)
Inter-methods correlation					
r	-0.01	0.89	0.65	0.24	0.01
[95% CI]	[-0.18, 0.15]	[0.82, 0.93]	[0.42, 0.79]	[-0.71, 0.88]	[-0.32, 0.35]
P-value	$P=0.86$	$P<0.001$	$P<0.001$	$P=0.64$	$P=0.94$
Bland-Altman analysis					
Absolute bias (mean±SD)	21.39 ± 48.02	-6.13 ± 6.69	5.27 ± 22.12	43.17 ± 22.84	86.37 ± 54.20
Limits of agreement	-72.72 to 115.5	-19.25 to 6.99	-38.08 to 48.62	-1.59 to 87.92	-19.87 to 192.61

r correlation coefficient, *CI* confidence interval, *SD* standard deviation

Recently, Brooks et al. introduced a 3D method of measuring femoral version [28], which has certain similarities to our approach. Specifically, they separately oriented the 3D models of the proximal and distal ends of the femur so that they were aligned along the long axis of the femur to enable femoral version measurements. This is conceptually similar to ours. However, since only images of the proximal and distal ends of the femur were available, they applied least-squares regression to separately align the two ends of the femur, and automate the calculation of femoral version. In comparison, we acquired CT images along the entire femur, and thus were able to generate a virtual femur model, which can be manipulated in 3D via PACS to measure femoral version during customary review of the study. As additional points of departure from our study, Brooks et al. compared their method to the 2D technique described by Reikeras [2], and their study sample consisted of only 17 children with limited pathologies (compared to 71 children in our study which are characterized by an assortment of deformities). The wide range of femoral versions that are present within our study sample further highlights the importance of having a generalizable femoral version assessment technique.

The imaging FOV for most 2D axial CT methods of measuring femoral version included only the hips and the knees, purposely excluding the mid femoral shafts to reduce radiation exposure. By design, our 3D femur model method required CT images of the entire femur to enable virtual rotation and spatial alignment of this bone. Naturally, this larger imaging FOV is associated with slightly higher radiation exposure than if only the proximal and distal femur were imaged. However, our studies were performed using low-radiation dose CT technique and the added radiation exposure is to the mid femoral shaft, an anatomical region that is of low radiation sensitivity. Consequently, any potential additional risk to the patient is expected to be minimal. In the future, we intend to quantitatively assess the excess radiation associated with our virtual 3D femur model method as compared to the conventional technique of just imaging the ends of the femurs.

Our retrospective study has a few limitations. Firstly, this study only compared our virtual 3D femur model method to the widely-accepted 2D axial slice method proposed by Murphy et al [1]. As most of the other proposed techniques in the literature are derivatives of the Murphy's method, we suspect that comparisons to other measurement techniques made off of 2D images would yield similar results. Secondly, the lack of ground truth femoral version measurements limited our ability to assess the accuracy of our novel methodology. Instead, we performed a reliability and agreement analysis, which is the standard statistical approach when

accuracy analysis is not possible [44, 48]. To further support our findings, we investigated the underlying cause for measurement variability through regression and subgroup analysis. Thirdly, by separating our study sample into 4 groups for subgroup analysis, our granular analysis had one subgroup consisting of only six participants (i.e. femurs with only hip flexion deformity). Caution should be exercised when interpreting the findings based on such a small sample. Lastly, this reliability study was based on a modest size patient sample with measurements performed by two expert pediatric musculoskeletal radiologists experienced with the proposed technique. Utilizing a larger and more heterogeneous study sample while employing radiologists with more variable backgrounds and expertise would further generalize our results.

Clinically, our virtual 3D femur model technique is straightforward to implement into the clinical work flow and intuitive to use. The technology required to deploy our methodology is widely available. This technique is the preferred method of assessing femoral version and surgical planning by our orthopedic colleagues. This time-tested virtual 3D femur model has been in use in our department for the past 10 years. Regardless, external validation studies could further confirm the generalizability and reliability of our method.

Conclusion

The virtual 3D femur model method for measuring femoral version is more reliable and generalizable than the Murphy's 2D axial slice method for evaluating femurs with apparent coxa valga and flexion hip deformities. Our study results suggest that our novel methodology may allow more reliable assessment of femoral version in children with hip deformities, which represent an important and substantial subgroup of patients in our clinical practice.

Authors contributions A.T. conceived, supervised and supported the study. J.I.N. collated the data, performed the statistical analysis and drafted the initial manuscript and multimedia. S.D.B. and M.A.B. interpreted the images. D.Y.J. and D.A. conceived and implemented the 3D model. E.L. helped perform the statistical analysis. All authors reviewed and approved the final manuscript.

Declarations

References

1. Murphy SB, Simon SR, Kijewski PK et al (1987) Femoral anteversion. J Bone Joint Surg Am 69:1169–1176
2. Reikerås O, Bjerkreim I, Sortland O (1983) Fluoroscopy in measurement of femoral neck anteversion. Acta Radiol Diagn (Stockh) 24:81–83
3. Sankar WN, Novais E, Koueiter D et al (2018) Analysis of femoral version in patients undergoing periacetabular osteotomy for symptomatic acetabular dysplasia. J Am Acad Orthop Surg 26:545–551
4. Roth PB (1920) A note on abnormal torsion of the femoral shaft. Proc R Soc Med 13:237–241
5. Fabricant PD, Fields KG, Taylor SA et al (2015) The effect of femoral and acetabular version on clinical outcomes after arthroscopic femoroacetabular impingement surgery. J Bone Joint Surg Am 97:537–543
6. Mastel MS, El-Bakoury A, Parkar A et al (2020) Outcomes of femoral de-rotation osteotomy for treatment of femoroacetabular impingement in adults with decreased femoral anteversion. J Hip Preserv Surg 7:755–763
7. Lerch TD, Antioco T, Meier MK et al (2022) Combined abnormalities of femoral version and acetabular version and McKibbin Index in FAI patients evaluated for hip preservation surgery. J Hip Preserv Surg 9:67–77
8. Sutter R, Dietrich TJ, Zingg PO, Pfirrmann CWA (2012) Femoral antetorsion: comparing asymptomatic volunteers and patients with femoroacetabular impingement. Radiology 263:475–483
9. Schaver AL, Oshodi A, Glass NA et al (2022) Cam morphology is associated with increased femoral version: Findings from a collection of 1,321 cadaveric femurs. Arthroscopy 38:831–836
10. Davids JR, Marshall AD, Blocker ER et al (2003) Femoral anteversion in children with cerebral palsy. Assessment with two and three-dimensional computed tomography scans. J Bone Joint Surg Am 85:481–488
11. Abel MF, Sutherland DH, Wenger DR, Mubarak SJ (1994) Evaluation of CT scans and 3-D reformatted images for quantitative assessment of the hip. J Pediatr Orthop 14:48–53
12. Sugano N, Noble PC, Kamaric E (1998) A comparison of alternative methods of measuring femoral anteversion. J Comput Assist Tomogr 22:610–614
13. Tomczak RJ, Guenther KP, Rieber A et al (1997) MR imaging measurement of the femoral antetorsional angle as a new technique: comparison with CT in children and adults. AJR Am J Roentgenol 168:791–794
14. Takeuchi S, Goto H, Iguchi H et al (2019) Ultrasonographic assessment of femoral torsion angle based on tilting angles of femoral neck and condylar axis. Ultrasound Med Biol 45:1970–1976
15. Kuo TY, Skedros JG, Bloebaum RD (2003) Measurement of femoral anteversion by biplane radiography and computed tomography imaging: comparison with an anatomic reference. Invest Radiol 38:221–229
16. Schmaranzer F, Kallini JR, Miller PE et al (2020) The effect of modality and landmark selection on MRI and CT femoral torsion angles. Radiology 296:381–390
17. Jarrett DY, Oliveira AM, Zou KH et al (2010) Axial oblique CT to assess femoral anteversion. AJR Am J Roentgenol 194:1230–1233
18. Hermann KL, Egund N (1997) CT Measurement of Anteversion in the Femoral Neck. Acta Radiol 38:527–532

19. Schmaranzer F, Lerch TD, Siebenrock KA et al (2019) Differences in femoral torsion among various measurement methods increase in hips with excessive femoral torsion. Clin Orthop Relat Res 477:1073–1083

20. Kaiser P, Attal R, Kammerer M et al (2016) Significant differences in femoral torsion values depending on the CT measurement technique. Arch Orthop Trauma Surg 136:1259–1264

21. Lerch TD, Kaim T, Hanke MS et al (2023) Assessment of femoral retroversion on preoperative hip magnetic resonance imaging in patients with slipped capital femoral epiphysis: Theoretical implications for hip impingement risk estimation. J Child Orthop 17:116–125

22. Zhang Z, Ren N, Cheng H et al (2023) Periacetabular osteotomy for Tönnis grade 2 osteoarthritis secondary to hip dysplasia. Int Orthop. https://doi.org/10.1007/s00264-023-05795-w

23. Meier MK, Wagner M, Brunner A et al (2023) Can gadolinium contrast agents be replaced with saline for direct MR arthrography of the hip? Eur Radiol, A pilot study with arthroscopic comparison. https://doi.org/10.1007/s00330-023-09586-0

24. Zurmühle CA, Kuner V, McInnes J et al (2021) The crescent sign - a predictor of hip instability in magnetic resonance arthrography. J Hip Preserv Surg 8:164–171

25. Schmaranzer F, Kheterpal AB, Bredella MA (2021) Best practices: Hip femoroacetabular impingement. AJR Am J Roentgenol 216:585–598

26. Buly RL, Sosa BR, Poultsides LA et al (2018) Femoral derotation osteotomy in adults for version abnormalities. J Am Acad Orthop Surg 26:e416–e425

27. Lee YS, Oh SH, Seon JK et al (2007) 3D femoral neck anteversion measurements based on the posterior femoral plane in ORTHO-DOC® system. Med Biol Eng Comput 45:325–325

28. Brooks JT, Bomar JD, Jeffords ME et al (2021) Defining a new three-dimensional method for determining femoral torsional pathology in children. J Pediatr Orthop B 30:331–336

29. Reikerås O, Bjerkreim I, Kolbenstvedt A (1983) Anteversion of the acetabulum and femoral neck in normals and in patients with osteoarthritis of the hip. Acta Orthop Scand 54:18–23

30. Gittings DJ, Dattilo JR, Fryhofer G et al (2017) Treatment of hip flexion contractures with psoas recession through the middle window of the ilioinguinal approach. JBJS Essent Surg Tech 7:e25

31. Pinero JR, Goldstein RY, Culver S et al (2012) Hip flexion contracture and diminished functional outcomes in cerebral palsy. J Pediatr Orthop 32:600–604

32. Tönnis D, Heinecke A (1999) Current concepts review - acetabular and femoral anteversion. J Bone Joint Surg Am 81:1747–1770

33. Morvan J, Bouttier R, Mazieres B et al (2013) Relationship between hip dysplasia, pain, and osteoarthritis in a cohort of patients with hip symptoms. J Rheumatol 40:1583–1589

34. Dolan MM, Heyworth BE, Bedi A et al (2011) CT reveals a high incidence of osseous abnormalities in hips with labral tears. Clin Orthop Relat Res 469:831–838

35. Liu RW, Toogood P, Hart DE et al (2009) The effect of varus and valgus osteotomies on femoral version. J Pediatr Orthop 29:666–675

36. Morrell DS, Pearson JM, Sauser DD (2002) Progressive bone and joint abnormalities of the spine and lower extremities in cerebral palsy. Radiographics 22:257–268

37. Mathews SS, Jones MH, Sperling SC (1953) Hip derangements seen in cerebral palsied children. Am J Phys Med 32:213–221

38. Sauser DD, Hewes RC, Root L (1986) Hip changes in spastic cerebral palsy. AJR Am J Roentgenol 146:1219–1222

39. Kay RM, Jaki KA, Skaggs DL (2000) The effect of femoral rotation on the projected femoral neck-shaft angle. J Pediatr Orthop 20:736–739

40. Boese CK, Jostmeier J, Oppermann J et al (2016) The neck shaft angle: CT reference values of 800 adult hips. Skeletal Radiol 45:455–463

41. Sullivan ES, Jones C, Miller SD et al (2022) Effect of positioning error on the Hilgenreiner epiphyseal angle and the head-shaft angle compared to the femoral neck-shaft angle in children with cerebral palsy. J Pediatr Orthop B 31:160–168

42. Sullivan ES, Jones C, Miller SD et al (2022) Effect of positioning error on the Hilgenreiner epiphyseal angle and the head-shaft angle compared to the femoral neck-shaft angle in children with cerebral palsy. J Pediatr Orthop B 31:160–168

43. Koo TK, Li MY (2016) A guideline of selecting and reporting intraclass correlation coefficients for reliability research. J Chiropr Med 15:155–163

44. Martin Bland J, Altman D (1986) Statistical methods for assessing agreement between two methods of clinical measurement. Lancet 327:307–310

45. Liljequist D, Elfving B, Skavberg Roaldsen K (2019) Intraclass correlation - A discussion and demonstration of basic features. PLoS One 14:e0219854

46. Cicchetti DV (1994) Guidelines, criteria, and rules of thumb for evaluating normed and standardized assessment instruments in psychology. Psychol Assess 6:284–290

47. Haghayegh S, Kang H-A, Khoshnevis S et al (2020) A comprehensive guideline for Bland-Altman and intra class correlation calculations to properly compare two methods of measurement and interpret findings. Physiol Meas 41:055012

48. Anvari A, Halpern EF, Samir AE (2018) Essentials of statistical methods for assessing reliability and agreement in quantitative imaging. Acad Radiol 25:391–396