



OPEN

Reliability of the pelvis and femur anatomical landmarks and geometry with the EOS system before and after total hip arthroplasty

Xavier Gasparutto^{1✉}, Pauline Besonhe², Peter Luca DiGiovanni², Matthieu Zingg², Sana Boudabbous³, Stéphane Armand¹ & Didier Hannouche²

Bi-plane X-ray provides 3D measurements of the lower limb based on the identification of anatomical landmarks in sagittal and frontal X-rays. In clinical practice, such measurements involve multiple operators and sessions. This study aimed at evaluating the reliability of anatomical landmarks identification and geometric parameters of the pelvis and femur measured with bi-plane X-rays before and after total hip arthroplasty (THA). Twenty-eight patients undergoing primary THA were selected retrospectively. Two operators performed three reconstructions for each patient before and after THA. Intraclass correlation (ICC) and smallest detectable change (SDC) were computed for intra-operator, inter-operator, and test–retest conditions. Most anatomical landmark positions had good to excellent SDC (< 5 mm) apart from the centre of the sacral slope, greater trochanter, and anterior superior iliac spines (up to 7.1, 16.9, and 21.5 mm respectively). Geometric parameters had moderate to excellent SDC, apart from femoral and stem torsion, pelvic incidence, and APP inclination with poor SDC (9–12°). The sagittal view had significantly higher measurement errors than the frontal view. Test–retest and inter-operator conditions had no significant differences suggesting a low influence of patient posture. Osteoarthritis and the presence of implants did not seem to influence reliability and measurement error. This study could be used as a reference when assessing lower limb structure with bi-plane X-rays.

Bi-plane X-ray is a low dose imaging system that provides 3D parameters of the spine and lower limb based on two X-ray images that are synchronised, spatially calibrated, and orthogonal (frontal and sagittal view). The 3D parameters are computed with four steps: (1) anatomical landmarks are manually identified on the images by a trained operator (e.g. acetabulum), (2) a statistical model is fitted to those landmarks, (3) the model is refined manually, and (4) the software computes predefined parameters based on the 3D geometry of the final model. This imaging method and associated software have been mainly designed to study global sagittal spinal balance, scoliosis and lower limb measurements for total hip arthroplasty (THA) (EOS imaging, Paris, France).

In the context of THA, measurements with bi-plane X-rays typically include measurements of multiple radiological parameters such as pelvic incidence, cup orientation, or femoral offset. These measurements involve multiple operators and sessions, i.e. measurements before and after surgery, with the identifications of anatomical landmarks on the pelvis, femur, and tibia. Quantifying the reliability of measurements is critical in clinical settings or when performing research studies on large databases. Indeed, the analysis of the results will be done according to this reliability. Reliability was defined by the COSMIN taxonomy¹ as the “degree to which the measurement is free from measurement error” and “the extent to which scores for patients who have not changed are the same for repeated measurement under several conditions”, e.g. multiple operators or multiple sessions. Since the present

¹Laboratoire de Cinésiologie Willy Taillard, Geneva University Hospitals and University of Geneva, Rue Gabrielle-Perret-Gentil 4, 1211 Geneva 14, Switzerland. ²Division of Orthopaedic Surgery and Musculoskeletal Trauma Care, Surgery Department, Geneva University Hospitals and University of Geneva, Geneva, Switzerland. ³Division of Radiology, Diagnosis Department, Geneva University Hospitals and University of Geneva, Geneva, Switzerland. ✉email: xavier.gasparutto@hcuge.ch

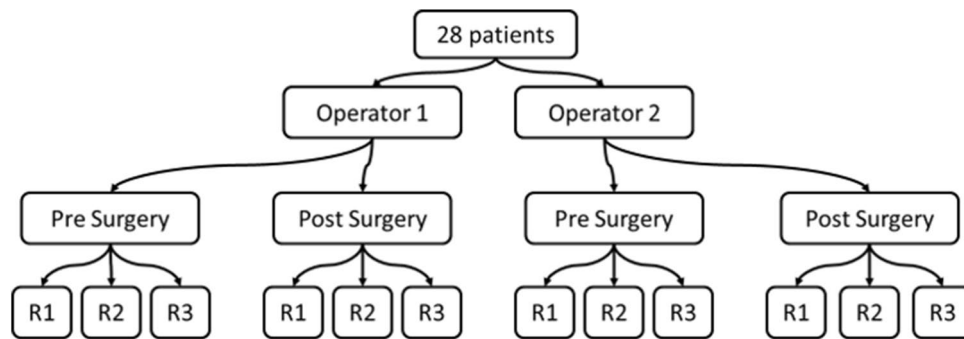


Figure 1. Measurement plan of the study.

study will focus on measurements with bi-plane X-rays and not on the definition of a specific clinical outcome, there are two relevant properties of the reliability domain: the reliability (proportion of total variance) and the measurement error (systematic and random error). For clarity, the reliability property of the reliability domain will be referred to as reliability in the following sections.

Several studies have evaluated the reliability domain of the lower limb parameters measured with bi-plane X-rays with mainly good to excellent reliability. Measurements of pelvic parameters (pelvic incidence, sacral slope, pelvic tilt) have shown excellent intra-operator and inter-operator reliability in adolescents with idiopathic scoliosis (intraclass correlation coefficient [ICC] > 0.9)² and good to excellent ICC (ICC > 0.795) for gross pelvic and acetabular morphology in one dry pelvis³. The measurement error of pelvis axial rotation was found to be lower than 1° on one dry pelvis⁴ and a decrease in intra-operator reliability was linked to an increase in axial rotation in six patients⁵. Regarding the acetabular cup orientation, the intra- and inter-operator reliabilities were found to be excellent (ICC > 0.9) with a confidence interval lower than 5° in ten patients with THA selected from a population of 44 patients⁶. The intra and inter-operator measurement errors of the cup inclination and anteversion were lower than 2.5° on one dry pelvis⁷. Finally, lower limb measurements (3D geometry of femur and tibia) were found to have good to excellent intra and inter-observer reliability (ICC > 0.75) in 25 patients planned for THA⁸ and lower limb length showed excellent reliability (ICC > 0.90) with one phantom limb⁹. Likewise, femoral torsion showed good to excellent inter-operator reliability (ICC > 0.86) in 30 patients with planned or completed THA¹⁰ and in cadaveric femur¹¹ while being slightly lower for a different cohort of patients with hip osteo-arthritis¹² (ICC of 0.711–0.769). Finally, the femoral offset showed good intra and inter-operator reliability (ICC > 0.75) in 110 patients with THA¹³.

These studies show encouraging results, but some limitations remain present in the context of THA. There is a lack of knowledge about the reliability of pelvic parameters for patients with THA. Moreover, ten patients seem like a small population to identify the intra and inter-operator reliability of cup orientation. Additionally, although phantoms and dry bones can be relevant for proofs of concept, the visibility of anatomical landmarks on such images is considerably higher than in clinical practice with THA patients. Such studies will likely provide the best-case scenario but it is expected that acquisitions performed in clinical practice will lead to lower levels of reliability and higher measurement errors. Furthermore, to the best of our knowledge, there is currently no information on test–retest reliability with patients or on the reliability of anatomical landmark identification, the only manual step performed by the operator.

Most importantly, most studies focus on reliability with ICC but a clear evaluation of measurement errors seems to be missing. To our knowledge, there is no report of smallest detectable changes (SDC) for the various measurements of positions, lengths, and angles performed with such technology.

Thus, this study aimed at evaluating the multiple properties of the reliability domain for the anatomical landmarks and geometry of the pelvis and femur measured with bi-plane X-rays in patients with THA. Tibia parameters were not included since they are not directly affected by the surgery.

Methods

Design of experiment. Two operators performed three reconstructions at each session for each patient (Fig. 1) with a research version of the sterEOS software (EOS imaging, Paris, France). Twenty-eight patients were selected to allow the computations of ICC as low as 0.3 with a power of 80%¹⁴. Indeed, Guenoun et al.⁸ found ICC as low as 0.443. The reconstructions of both sessions were performed simultaneously and the delay between each reconstruction for a patient was 20 days. The operators were one trained researcher specialized in movement analysis with two years of experience in 3D reconstruction of the lower limb and one orthopaedic surgeon trained for 3D reconstructions and with extensive experience in the analysis of X-rays. This design has four degrees of freedom, i.e. sources of variance: the operators, the visits, the reconstructions, and the patients.

Population. Twenty-eight patients were randomly selected retrospectively from a database of 439 patients with end-stage primary osteoarthritis that underwent primary unilateral THA at the Geneva University Hospitals between 2016 and 2019 and that had a bi-plane X-ray before and two months after surgery. Random selection of patients was constrained to avoid differences in distribution (Kolmogorov–Smirnov test) and mean (Student’s t-test) for age, height, and weight, and in sex distribution (χ^2 test) (Table 1). This study was approved

	Database	Subset	p-value	
			Student's	KS-test
N patients	439	28	–	–
Age (years)	66.6 (12.7)	65.7 (11.7)	0.714	0.899
Height (cm)	166.8 (9.4)	168 (9.3)	0.521	0.992
Weight (kg)	76.2 (16.6)	77.8 (17.5)	0.634	0.960
Sex (% female)	56.70%	60.70%	0.828	
Side OA (% left)	46%	46.40%	1.000	

Table 1. Patient characteristics. Age, height, weight are described as: mean (SD). Age, height and weight differences and distributions were tested with unpaired Student's t-test and Kolmogorov-Smirnoff (KS) test, and sex and side of osteoarthritis (OA) distributions were tested with χ^2 .

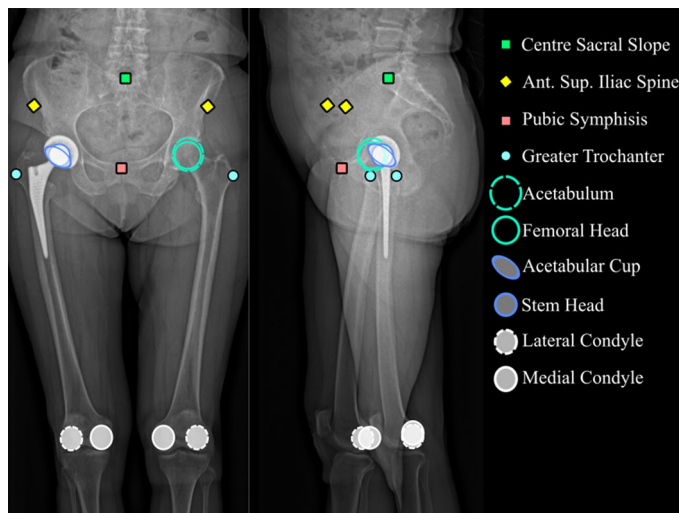


Figure 2. Anatomical landmarks. *Ant. Sup* anterior superior.

by the local ethics committee (CCER Geneva, Switzerland). Patient consent was obtained and protected by the Geneva Arthroplasty Registry and all experiments were performed in accordance with relevant guidelines and regulations. Only one acquisition was done per session (before and after surgery) to avoid increasing the radiation dose exposure of the patients.

Anatomical landmarks. The considered anatomical landmarks (Fig. 2) were the points used to define the geometric model.

Pelvis and femur geometry. The selected 3D geometric parameters of the pelvis and femur were the standard output of sterEOS (Fig. 3). To assess reliability between sessions, parameters that did not depend on posture and that were not affected by surgery were selected (Fig. 4).

Statistical analysis. According to the COSMIN checklist¹, reliability was assessed using ICC and measurement error with SDC. The ICC is defined as the correlations between two measurements that have exactly the same components except for the 'class' component¹⁵. Variance of class components was computed from a single measure, two-way mixed effects model^{16,17} with the lme4 package¹⁸ of R¹⁹. The total variance was computed as the sum of the variance of class components:

$$\sigma_{Total}^2 = \sigma_p^2 + \sigma_{po}^2 + \sigma_{pv}^2 + \sigma_{pm}^2 + \sigma_{residual}^2$$

where each term refers to a variance component: p = patient, o = operator, v = visit, m = repeated measure.

Intra-rater, inter-rater and test-retest ICCs were computed using the following formula²⁰:

Intra-rater reliability:

$$ICC_{Intra} = \frac{\sigma_{Total}^2 - (\sigma_{pm}^2 + \sigma_{residual}^2)}{\sigma_{Total}^2}$$

Inter-rater reliability:

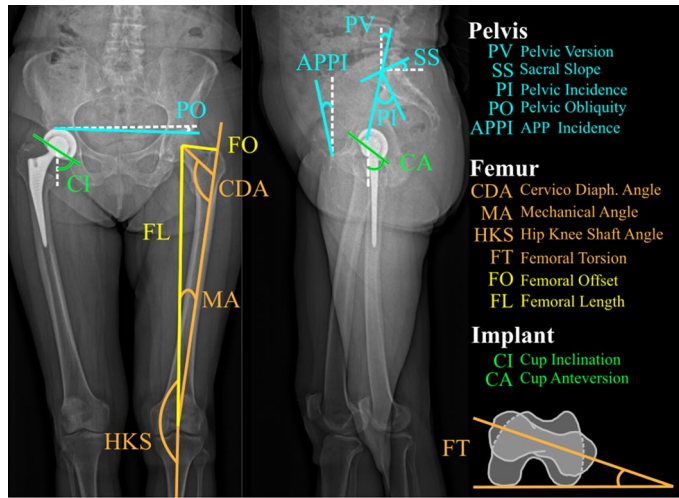


Figure 3. Pelvis and femur geometry parameters. *Diaph* diaphyseal.

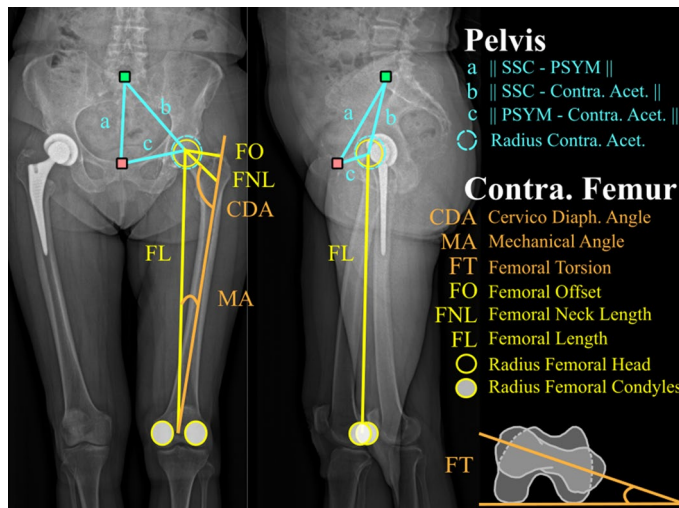


Figure 4. Test-retest parameters. SSC centre of sacral slope, PSYM pubic symphysis, *Contra* contralateral, *Acet* acetabulum centre, *Diaph* diaphyseal.

$$ICC_{inter} = \frac{\sigma_{Total}^2 - (\sigma_{po}^2 + \sigma_{residual}^2)}{\sigma_{Total}^2}$$

Test-retest reliability:

$$ICC_{test-retest} = \frac{\sigma_{Total}^2 - (\sigma_{pv}^2 + \sigma_{residual}^2)}{\sigma_{Total}^2}$$

The standard error of measurement was then calculated as $SEM = \sqrt{\sigma^2 * (1 - \rho)}$ ¹⁶ where σ^2 represents the total variance in the ICC formula, and ρ is the ICC. Finally, the SDC was calculated as $SDC = 1.96 * \sqrt{2} * SEM$ ¹⁶.

The ICC were classified as: poor (<0.5), moderate (0.5–0.75), good (0.75–0.9), and excellent (>0.9)²¹. Victor et al.²² found a standard deviation around 1 mm and 1 degree for intra and inter-operator measurement error in femoral landmarks and anatomical axes identified with CT scan, the gold-standard method. This is equivalent to a SDC of 2.8 mm/° which led us to define the following classification of SDC levels: poor (> 10 mm/°), moderate (5–10 mm/°), good (3–5 mm/°), excellent (<3 mm/°).

To refine the understanding of the anatomical landmarks’ measurement errors, paired Student’s t-tests ($p < 0.05$) were performed between the SDC obtained in different directions (antero-posterior, medial-lateral, vertical). Welch two-sample t-tests were performed between directions and the radiuses ($p < 0.05$). To evaluate the consistency between different conditions, Student’s t-tests were performed between intra-operator, inter-operator,

	Intra-operator				Inter-operator			
	AP	ML	V	R	AP	ML	V	R
Pelvis								
Pubic Symphysis	5.0	1.8	3.2	-	5.5	1.8	3.4	-
Centre Sacral Slope	6.3	5.9	4.9	-	7.1	5.9	5.0	-
Left Ant. Sup. Iliac Spine	15.3	6.2	9.5	-	21.5	9.6	11.8	-
Right Ant. Sup. Iliac Spine	17.2	8.3	14.4	-	19.1	8.9	16.0	-
Homo. Acetabulum	3.6	3.2	3.0	2.4	3.7	4.9	2.8	5.4
Contra. Acetabulum	4.8	2.5	2.6	2.3	4.8	3.4	4.2	4.6
Contra. Acet. pre-THA	4.1	4.1	2.3	2.0	4.8	3.2	3.6	5.0
Contra. Acet. post-THA	5.0	2.7	2.6	2.4	5.1	3.5	4.8	4.1
Femur								
Homo. Fem. Head pre-THA	4.7	1.9	2.5	1.9	4.5	2.5	2.8	2.3
Contra. Fem. Head	4.6	1.2	1.5	1.4	4.7	1.4	1.7	1.5
Contra. Fem. Head pre-THA	4.6	1.2	1.4	1.4	4.7	1.4	1.6	1.4
Contra. Fem. Head post-THA	4.6	1.2	1.5	1.4	4.8	1.6	1.8	1.6
Left Gr. Troch. pre-THA	6.0	4.4	2.6	-	5.2	5.0	2.9	-
Right Gr. Troch. pre-THA	5.2	4.5	2.8	-	6.0	4.5	2.9	-
Contra. Gr. Troch. post-THA	15.1	4.2	2.6	-	16.9	4.4	2.4	-
Contra. Lateral Condyle	2.7	3.2	1.2	1.8	2.9	3.7	1.5	2.0
Contra. Medial Condyle	3.2	2.8	1.3	1.8	3.8	4.1	1.5	2.6
Homo. Lateral Condyle	3.2	4.0	1.3	1.7	3.6	4.3	1.4	1.8
Homo. Medial Condyle	4.7	3.1	1.3	1.7	4.9	3.5	1.3	2.4
Implant								
Acetabular Cup	1.4	1.5	1.5	1.2	1.7	1.6	1.5	1.7
Stem Femoral Head	0.8	0.8	0.8	0.7	0.7	0.8	0.9	0.7
Legend								
Excellent (<3mm)		Good (3 to 5 mm)		Moderate (5 to 10mm)		Poor (>10mm)		

Table 2. Intra and inter-operator smallest detectable change (SDC) of the position of anatomical points (in mm). *AP* anterior–posterior direction, *ML* medio-lateral direction, *V* vertical direction, *R* radius, *Ant* anterior, *Sup* superior, *Homo* homolateral side, *Contra* contralateral side, *Acet* acetabulum, *Fem* femoral, *THA* total hip arthroplasty, *Gr. Troch* Greater Trochanter.

and test–retest SDC obtained for parameters independent of posture and surgery. Additionally, Student's t-tests were performed between parameters of the right and left lower limbs pre-surgery to assess potential systematic bias since the lateral source of radiation is on the left-hand side of the patient.

Results

Values of parameters for all patients are reported in supplementary files S1–S8. Two patients (P03 and P09) presented inversion of condyles on the sagittal view and were excluded from the analysis of the related parameters. Results including these two patients are reported in Supplementary file S9. A graphical overview of the results is presented in Supplementary file S10.

Anatomical landmarks. The SDCs were good to excellent for most anatomical landmarks apart from the centre of the sacral slope (good to moderate), the anterior superior iliac spines (moderate to poor), and the greater trochanter (moderate to poor) (Table 2).

The anterior–posterior direction had significantly higher SDC than medio-lateral ($p_{\text{intra}} = 0.002$, $p_{\text{inter}} = 0.006$), vertical ($p_{\text{intra}} < 0.001$, $p_{\text{inter}} < 0.001$) and radiuses ($p_{\text{intra}} < 0.001$, $p_{\text{inter}} = 0.006$) in both intra and inter-operator conditions. Medio-lateral intra-operator SDC was significantly higher than radiuses ($p_{\text{intra}} = 0.002$). Other differences were not significant.

The intra and inter-operator reliability of the position of all anatomical landmarks was excellent (Table 3). The intra-operator reliability of the radius of the condyles was good but the contralateral medial condyle radius had moderate reliability. The radius of the acetabulum had poor inter-operator reliability.

Pelvis and femur geometry. Most parameters had good to excellent intra and inter-operator reliability (Tables 4 and 5), apart from APP inclination (moderate inter), femoral torsion and cup inclination (moderate intra and inter), and the diameter of the acetabulum (poor inter). Test–retest reliabilities (Table 4) were good to

	Intra-operator				Inter-operator			
	AP	ML	V	R	AP	ML	V	R
Pelvis								
Pubic Symphysis	0.997	0.994	1.000	-	0.997	0.994	1.000	-
Centre Sacral Slope	0.996	0.953	0.999	-	0.995	0.953	0.999	-
Left Ant. Sup. Iliac Spine	0.979	0.976	0.997	-	0.959	0.940	0.995	-
Right Ant. Sup. Iliac Spine	0.973	0.975	0.992	-	0.967	0.972	0.990	-
Homolateral Acetabulum	0.999	0.996	1.000	0.879	0.998	0.990	1.000	0.354
Contralateral Acetabulum	0.998	0.995	1.000	0.860	0.998	0.991	0.999	0.468
Contra. Acet. pre-THA	0.998	0.998	1.000	0.900	0.997	0.993	0.999	0.365
Contra. Acet. post-THA	0.997	0.993	1.000	0.841	0.997	0.988	0.999	0.556
Femur								
Homo. Fem. Head pre-THA	0.998	0.998	1.000	0.909	0.998	0.997	1.000	0.872
Contra. Fem. Head	0.998	0.999	1.000	0.937	0.998	0.998	1.000	0.929
Contra. Fem. Head Pre-THA	0.998	0.999	1.000	0.942	0.998	0.999	1.000	0.937
Contra. Fem. Head Post-THA	0.998	0.999	1.000	0.939	0.998	0.998	1.000	0.913
Left Gr. Troch. pre-THA	0.997	0.994	1.000	-	0.998	0.992	1.000	-
Right Gr. Troch. pre-THA	0.997	0.990	1.000	-	0.995	0.990	1.000	-
Contra. Gr. Troch. post-THA	0.979	0.986	1.000	-	0.974	0.984	1.000	-
Contra. Lateral Condyle	0.999	0.995	1.000	0.861	0.999	0.993	1.000	0.822
Contra. Medial Condyle	0.999	0.996	1.000	0.884	0.998	0.992	1.000	0.740
Homo. Lateral Condyle	0.999	0.993	1.000	0.828	0.999	0.992	1.000	0.817
Homo. Medial Condyle	0.998	0.995	1.000	0.875	0.997	0.994	1.000	0.744
Implant								
Acetabular Cup	1.000	0.998	1.000	0.929	1.000	0.998	1.000	0.851
Stem Femoral Head	1.000	1.000	1.000	0.955	1.000	1.000	1.000	0.950
Legend								
Excellent (>0.9)	Good (0.75 to 0.9)			Moderate (0.5 to 0.75)	Poor (<0.5)			

Table 3. Intra and inter-operator intra-class correlation (ICC) of the position of anatomical points. *AP* anterior–posterior direction, *ML* medio-lateral direction, *V* vertical direction, *R* radius, *Ant* anterior, *Sup* superior, *Homo* homolateral side, *Contra* contralateral side, *Acet* acetabulum, *Fem* femoral, *THA* total hip arthroplasty, *Gr. Troch* Greater Trochanter.

excellent apart from the cervico-diaphyseal angle, the mechanical angle of the femur, and the femoral torsion (moderate).

The SDCs were moderate to excellent for most parameters but the femoral torsion post-THA, and pelvic incidence were poor for both inter and intra-operator conditions while stem antetorsion, cup anteversion with respect to the APP, and APP inclination were poor for the inter-operator condition only. Test–retest SDCs were moderate to excellent apart from the femoral torsion (poor).

Consistency between conditions. The intra-operator SDCs and ICCs were significantly lower than the test–retest condition ($p < 0.001$ and $p = 0.001$ respectively). There were no significant differences in ICCs or SDCs between the inter-operator and test–retest conditions and between the right and left sides (Supplementary file S11).

Discussion

The aim of this study was to evaluate the reliability and measurement error of the anatomical landmarks and geometry of the pelvis and femur evaluated with bi-plane X-rays in patients before and after THA.

The intra and inter-operator reliabilities were similar to the literature with most ICC above 0.75^{8,23,24}, apart from the inter-operator acetabulum radius. This consistency with the literature shows that the level of training of the operators met the standard for such measurements.

The poor inter-operator ICC for the acetabulum radius indicates a slightly different practice between operators for this parameter since intra-operator ICC was good. This indicates that the variability between patients was low, leading to a poor ICC despite a good SDC. Conversely, the pelvic incidence and antero-posterior position of the anterior–superior iliac spines had poor SDC but excellent ICC. This indicates a high variability between patients leading to an excellent ICC despite the poor SDC. Those two results underline the importance of reporting both ICC and SDC for a full understanding of the reliability domain.

	ICC			SDC		
	Intra	Inter	TRtest	Intra	Inter	TRtest
Pelvis						
Diameter Acet. Contralateral (mm)	0.860	0.468	0.843	2.3	4.6	2.5
Distance Pubic Symp. to Contra. Acet. (mm)	0.946	0.941	0.887	4.1	4.3	5.9
Distance Centre Sacral Slope to Contra. Acet. (mm)	0.932	0.894	0.916	6.4	8.0	7.1
Distance Centre Sacral Slope to Pubic Symp. (mm)	0.966	0.957	0.949	5.5	6.2	6.7
Femur						
Femur Length Contralateral (mm)	0.999	0.999	0.998	2.4	2.5	2.6
Femur Head Diameter Contralateral (mm)	0.969	0.958	0.965	2.0	2.3	2.1
Femoral Neck Length Contralateral (mm)	0.950	0.947	0.911	3.2	3.3	4.3
Femoral Offset Contralateral (mm)	0.916	0.915	0.853	4.3	4.3	5.7
Hip Knee Shaft Angle Contralateral. (°)	0.904	0.879	0.801	1.1	1.2	1.6
Cervico Diaphyseal Angle Contralateral (°)	0.784	0.797	0.734	5.8	5.7	6.5
Mechanical Angle Contralateral (°)	0.826	0.812	0.766	2.2	2.3	2.5
Femoral Torsion Contralateral (°)	0.789	0.829	0.769	10.7	9.6	11.1
Legend						
Excellent	Good	Moderate	Poor			

Table 4. Intra-operator (Intra), intra-operator (Inter) and test–retest (TRtest) ICC and SDC of pelvis and femur geometrical parameters independent of posture and surgery. *Symp* symphysis, *Contra* contralateral, *Acet* acetabulum.

	ICC		SDC	
	Intra	Inter	Intra	Inter
Pelvis				
Axial Rotation (°)	0.992	0.991	1.9	1.9
Obliquity (°)	0.968	0.928	3.4	5.1
Sacral Slope (°)	0.896	0.904	9.8	9.4
Pelvic Version (°)	0.950	0.936	4.0	4.5
Pelvic Incidence (°)	0.894	0.900	11.5	11.2
APP Inclination (°)	0.809	0.658	7.7	10.4
Femur				
Femur Length Homolateral (mm)	0.998	0.998	2.7	2.8
Femoral Offset Homo. Pre-THA (mm)	0.926	0.912	4.1	4.5
Femoral Offset Homo. Post-THA (mm)	0.981	0.981	2.0	2.0
Fem. Neck Length Homo. Pre-THA (mm)	0.943	0.925	3.3	3.8
Hip Knee Shaft Angle Homolateral (°)	0.904	0.890	1.4	1.5
Cervico Diaph. Angle Homolateral (°)	0.836	0.826	5.7	5.9
Femoral Torsion Contra. Pre-THA (°)	0.862	0.834	8.5	9.4
Femoral Torsion Contra. Post-THA (°)	0.768	0.781	11.3	11.0
Implant				
Cup Anteversion w.r.t APP (°)	0.860	0.774	9.9	12.5
Cup Anteversion w.r.t cabin (°)	0.873	0.852	8.2	8.9
Cup Inclination w.r.t APP (°)	0.725	0.603	6.0	7.2
Cup Inclination w.r.t cabin (°)	0.884	0.866	3.8	4.0
Antetorsion Stem (°)	0.918	0.865	8.9	11.4
Legend				
Excellent	Good	Moderate	Poor	

Table 5. Intra-operator (Intra) and inter-operator (Inter) ICC and SDC of pelvis, femur and implant geometrical parameters dependent of posture and/or surgery. *APP* anterior pelvic plane, *Homo* homolateral side, *Fem* femoral, *THA* total hip arthroplasty, *Diaph* diaphyseal, *Contra* contralateral, *w.r.t* with respect to.

Test–retest reliability and measurement error were assessed for the first time with patients and were moderate to excellent, apart from the femoral torsion (poor). They were rated as good as the inter-operator condition, thus, the change in posture of patients between measurements does not seem to have a strong influence on parameters independent of posture, as suggested previously^{25,26}. This underlines that bi-plane X-rays are relevant to compare pre- and post-THA measurements.

Likewise, during bi-plane X-rays, the position of the patient in the cabin is standardised and the left lower limb is closer to the source of radiation. This results in a larger magnification of the left side compared to the right side on the images which is often used during reconstruction. There were no differences between parameters assessed on the left and right lower limb which indicates the absence of systematic bias on the measurement error and reliability associated with the side.

The SDC and ICC of the contralateral acetabulum and contralateral femoral head were similar before and after THA. This suggests that the presence of the hip implant does not impede the quality of the reconstruction, apart from cases where the implant would completely overlap with the contralateral hip, and strengthens the validity of the test–retest evaluation presented in this study. To avoid overlapping issues, patients were positioned with one foot slightly forward as recommended²⁷. Still, the influence of the implant might not be negligible for all parameters, as seen by the slight increase in SDC for contralateral femoral torsion post-THA and especially with the 10 mm increase in SDC of the greater trochanter anterior–posterior position.

Likewise, there was no difference between the contralateral and homolateral acetabulum and femoral head ICC and SDC before surgery. This suggests that the measurements remain reliable with the presence of coxarthrosis.

The position of the stem femoral head had the best SDC which resulted in higher reliability of the implant's femoral offset when compared to the native hip.

Anterior superior iliac spines were the least reliable points followed by the centre of the sacral slope indicating that pelvic anatomical landmarks are more difficult to identify than femoral landmarks. These results were expected since the visibility on the pelvis can be low, especially on the sagittal view, due to the superposition of pelvic bone structures, and/or low bone density seen in elderly patients and inherent to this technology. Since the pubic symphysis had good to excellent intra-operator SDC and moderate to excellent inter-operator SDC, the identification of the anterior superior iliac spines seems to be responsible for the difficulties in defining the APP, also reported by Tokunaga et al.²⁸.

Positions in the anterior–posterior direction showed significantly higher SDC than in the other two directions. This indicates that the anatomical landmark identification on the sagittal view is more critical than on the frontal view. Errors on this view could be responsible for the poor SDC obtained for APP inclination, pelvic incidence, femoral torsion, and stem antetorsion. Indeed, the normal to the sacral slope, used to compute pelvic incidence, is identified on this view. The poor SDC measured for pelvic incidence was consistent with a previous study that found within-subject standard deviation of 4.12°²³. Since these parameters characterised the standard error of measurements, it would be equivalent to a within-subject SDC of 11.6°, consistent with the intra- and inter-operator SDC of 11.5° and 11.2° found in the present study. The sacral slope was identified as the main factor for this variability on sagittal X-rays²⁹ which is consistent with the SDC of 9.8° and 9.4° measured for this parameter.

Femoral torsion and stem antetorsion had moderate to poor SDC (8.9 to 11.4°). These SDC could be linked to the moderate to poor SDC in the anterior–posterior direction of the greater trochanter. Indeed, the SDC of the femoral torsion was slightly higher after surgery while the SDC of the greater trochanter increased by 10 mm post-surgery. Another candidate could be the anterior–posterior position of the condyles. Although those SDC are good, these points are close to each other in the medial–lateral direction (mean of 48.5 ± 3.7 mm in the present study) and a small position error could lead to a significant angular error. Vanhove et al.³⁰ found mean differences of 1.3° between two radiologists for the measurements of femoral torsion with the CT scan based Weiner method. These results are equivalent to an SDC of 3.9° which is lower than our results. Comparisons between bi-plane X-ray measurements and CT scan showed differences of 1.3 ± 6.5° with 31% of patients over 5° of differences²⁸, of 4 ± 4°³¹, between – 5 and 7°³², and finally limits of agreements of – 5.0 to 4.2°³³ and – 17.53° to 10.82°¹². Interestingly, the limits of agreement presented by Cho et al.¹² have the same order of magnitude as the SDC found in the present study. Correlations between measurement of femoral torsion with CT scan and bi-plane X-rays were found to be strong ($R > 0.8$)^{10,34} but Mayr et al. showed that for patients with decreased (< 10°) and increased (> 20°) anteversion the correlation was only moderate to low, respectively³⁴. Although it was shown that CT scan was influenced by femur position, contrary to bi-plane X-rays^{26,34}, the higher measurement error with bi-plane X-rays tempers the recommendation of using this method over CT scan for accurate measurement of femoral torsion. Novel methods of automatic reconstruction of the femur with bi-plane X-rays might improve the reliability and measurement errors of this parameter³⁵.

Cup anteversion had moderate to poor SDC (8.2–12.5°). Cup anteversion measured with CT scan was found to have intra- and inter-operator measurement error of 0.52° and 3.2° respectively³⁶ corresponding to SDC of 1.5 and 9.0° respectively. Intra and inter-operator SDC with bi-plane X-rays in the present study were higher than SDC reported for CT scans³⁶ and seem to contradict a previous study showing equivalence between CT and bi-plane X-rays⁷. However, Journé et al.⁷ performed a comparison on one dry pelvis which represents an ideal case with much better visibility than the one obtained with patients. Comparison of cup inclination and anteversion measured in THA patients with EOS and CT scan showed mean differences below 3° in two studies^{28,31}. However, 17% and 34% of patients had more than 5° of difference for inclination and anteversion respectively, indicating that differences at the patient level can be large between CT and EOS measurements²⁸. Moreover, limits of agreement of – 5.4° to 6.6° were found in a third study between CT and EOS measurements³³. Still, correlations above 0.7 were found when comparing cup angles measured with CT and EOS²⁸.

Tokunaga et al.²⁸ found that stem antetorsion and cup anteversion had almost two times more patients above 5° of difference with CT scans than cup inclination. Interestingly, our results showed that these two first

parameters were found to have higher SDC than cup inclination. The larger measurement errors observed in our study could be linked to the larger discrepancy observed by Tokunaga et al.²⁸

Our study focused on reliability and measurement error, but high reliability does not necessarily translate to high validity. As presented in the previous sections, the validity of bi-plane X-rays has been assessed for multiple key parameters. Future studies should continue exploring the validity of this measurement device and compare its reliability and measurement errors to the ones of the gold-standard CT scans.

Despite warnings from the software, two patients presented inversion of condyles on the sagittal view. The error was identified by comparing the multiple reconstructions to each other. Those patients were excluded from the analysis, however, in clinical settings, this error could remain undetected and lead to a biased comparison of pre- and post-THA measurements. Keeping in mind the normal range of femoral torsion may help identify this issue³⁷. This outcome underlines the fact that the sagittal view seems to be the main source of error with bi-plane X-rays.

This study has some limitations. The number of patients could have been increased for a higher statistical power, but 80% seems reasonable. Only two operators were considered while the number of operators is frequently larger in a radiological department. Also, the time constraint being lower in research studies than in clinical practice, the present results might represent a best-case scenario. Our study only included patients with end-stage primary osteoarthritis and excluded patients with different pathologies such as severe hip dysplasia, post-traumatic arthritis or post-infection sequelae. Those pathologies can lead to severe deformities of the hip joint and may lead to higher measurement error and lower reliability. Indeed, the identification of structure may be more difficult and those deformities may reach the limits of the generic geometric statistical model used within the reconstruction software. Future studies should explore those limitations and could use the present methodology.

Conclusion

This study assessed the reliability and measurement error of anatomical landmark position and standard geometrical parameters for the pelvis and femur measured by bi-plane X-rays (EOS system). The sagittal view had significantly higher measurement errors than the frontal view. Osteoarthritis and the presence of an implant on the lateral view did not seem to have a strong influence on reliability and measurement errors. There were no significant differences between the test–retest and inter-operator reliability and measurement error which suggests that this technology is suitable to compare pre- and post-THA parameters.

Pelvic incidence and position of the anterior superior iliac spines had poor inter and intra-operator measurement errors that could be linked to higher errors on the sagittal view. The femoral and stem torsion, cup anteversion, and inclination had moderate to poor measurement errors. Interestingly, based on the literature, high reliability is challenging with CT scan for those parameters. Parameters with good SDC and poor ICC and vice-versa shows the importance of considering the whole domain of reliability to fully understand the measurement.

This study can help radiographers better understand their tool and define more accurate guidelines for parameters with lower reliability. It can also be used by clinicians when analysing bi-plane X-rays and in future longitudinal studies using this technology to assess the structure of the lower limb.

Data availability

Data is available on Yareta: <https://doi.org/10.26037/yareta:2a17pd6y7rgkzkd6nolwp6ucji>.

Received: 11 February 2022; Accepted: 7 December 2022

Published online: 11 December 2022

References

- Mokkink, L. B. *et al.* The COSMIN checklist for assessing the methodological quality of studies on measurement properties of health status measurement instruments: An international Delphi study. *Qual. Life Res.* **19**, 539–549. <https://doi.org/10.1007/s11136-010-9606-8> (2010).
- Ilharreborde, B., Ferrero, E., Alison, M. & Mazda, K. EOS microdose protocol for the radiological follow-up of adolescent idiopathic scoliosis. *Eur. Spine J.* **25**, 526–531. <https://doi.org/10.1007/s00586-015-3960-8> (2016).
- Bittershol, B. *et al.* EOS imaging of the human pelvis: Reliability, validity, and controlled comparison with radiography. *J. Bone Joint Surg.* **95**, 1–9 (2013).
- Rousseau, M.-A., Brusson, A. & Lazennec, J.-Y. Assessment of the axial rotation of the pelvis with the EOS® imaging system: Intra- and inter-observer reproducibility and accuracy study. *Eur. J. Orthop. Surg. Traumatol.* **24**, 891–895. <https://doi.org/10.1007/s00590-013-1281-3> (2014).
- Ghostine, B. *et al.* Influence of patient axial malpositioning on the trueness and precision of pelvic parameters obtained from 3D reconstructions based on biplanar radiographs. *Eur. Radiol.* **27**, 1295–1302 (2017).
- Barbier, O., Skalli, W., Mainard, L. & Mainard, D. The reliability of the anterior pelvic plane for computer navigated acetabular component placement during total hip arthroplasty: Prospective study with the EOS imaging system. *Orthop. Traumatol. Surg. Res.* **100**, S287–S291. <https://doi.org/10.1016/j.otsr.2014.07.003> (2014).
- Journé, A., Sadaka, J., Bécicourt, C. & Sautet, A. New method for measuring acetabular component positioning with EOS imaging: Feasibility study on dry bone. *Int. Orthop.* **36**, 2205–2209. <https://doi.org/10.1007/s00264-012-1650-2> (2012).
- Guenoun, B., Zedegan, F., Aim, F., Hannouche, D. & Nizard, R. Reliability of a new method for lower-extremity measurements based on stereoradiographic three-dimensional reconstruction. *Orthop. Traumatol. Surg. Res.* **98**, 506–513. <https://doi.org/10.1016/j.otsr.2012.03.014> (2012).
- Escott, B. G. *et al.* EOS low-dose radiography: A reliable and accurate upright assessment of lower-limb lengths. *J. Bone Joint Surg. Am.* **95**, e1831–1837. <https://doi.org/10.2106/jbjs.L.00989> (2013).
- Folainis, D. *et al.* Measuring femoral and rotational alignment: EOS system versus computed tomography. *Orthop. Traumatol. Surg. Res.* **99**, 509–516. <https://doi.org/10.1016/j.otsr.2012.12.023> (2013).

11. Pomerantz, M. L., Glaser, D., Doan, J., Kumar, S. & Edmonds, E. W. Three-dimensional biplanar radiography as a new means of accessing femoral version: A comparative study of EOS three-dimensional radiography versus computed tomography. *Skelet. Radiol.* **44**, 255–260. <https://doi.org/10.1007/s00256-014-2031-2> (2015).
12. Cho, B. W. *et al.* Evaluation of the reliability of lower extremity alignment measurements using EOS imaging system while standing in an even weight-bearing posture. *Sci. Rep.* **11**, 22039. <https://doi.org/10.1038/s41598-021-01646-z> (2021).
13. Lazennec, J. Y., Brusson, A., Dominique, F., Rousseau, M. A. & Pour, A. E. Offset and anteversion reconstruction after cemented and uncemented total hip arthroplasty: an evaluation with the low-dose EOS system comparing two- and three-dimensional imaging. *Int. Orthop.* **39**, 1259–1267. <https://doi.org/10.1007/s00264-014-2616-3> (2015).
14. Bujang, M. A. & Baharum, N. A simplified guide to determination of sample size requirements for estimating the value of intraclass correlation coefficient: A review. *Arch. Orofac. Sci.* **12**, 1–11 (2017).
15. Chia, K. & Sangeux, M. Quantifying sources of variability in gait analysis. *Gait Posture* **56**, 68–75. <https://doi.org/10.1016/j.gaitpost.2017.04.040> (2017).
16. De Vet, H. C., Terwee, C. B., Mokkink, L. B. & Knol, D. L. *Measurement in medicine: a practical guide.* (Cambridge university press, 2011).
17. McGraw, K. O. & Wong, S. P. Forming inferences about some intraclass correlation coefficients. *Psychol. Methods* **1**, 30 (1996).
18. Bates, D., Mächler, M., Bolker, B. & Walker, S. fitting linear mixed-effects models using lme4. *J. Stat. Softw.* **67**, 48. <https://doi.org/10.18637/jss.v067.i01> (2015).
19. R: A Language and Environment for Statistical Computing (R Foundation for Statistical Computing, 2021).
20. Van Lummel, R. C. *et al.* Intra-Rater, inter-rater and test-retest reliability of an instrumented timed up and go (iTUG) Test in patients with Parkinson's disease. *PLoS ONE* **11**, e0151881 (2016).
21. Koo, T. K. & Li, M. Y. A guideline of selecting and reporting intraclass correlation coefficients for reliability research. *J. Chiropr. Med.* **15**, 155–163. <https://doi.org/10.1016/j.jcm.2016.02.012> (2016).
22. Victor, J. *et al.* How precise can bony landmarks be determined on a CT scan of the knee?. *Knee* **16**, 358–365. <https://doi.org/10.1016/j.knee.2009.01.001> (2009).
23. Demzik, A. L. *et al.* Inter-rater and intra-rater repeatability and reliability of EOS 3-dimensional imaging analysis software. *J. Arthroplasty* **31**, 1091–1095. <https://doi.org/10.1016/j.arth.2015.11.026> (2016).
24. Clavé, A. *et al.* Reproducibility of length measurements of the lower limb by using EOS™. *Musculoskelet. Surg.* **102**, 165–171. <https://doi.org/10.1007/s12306-017-0518-4> (2018).
25. Melhem, E., Assi, A., El Rachkidi, R. & Ghanem, I. EOS(*) biplanar X-ray imaging: Concept, developments, benefits, and limitations. *J. Child. Orthop.* **10**, 1–14. <https://doi.org/10.1007/s11832-016-0713-0> (2016).
26. Morvan, G., Guerini, H., Carré, G. & Vuillemin, V. Femoral torsion: Impact of femur position on CT and stereoradiography measurements. *Am. J. Roentgenol.* **209**, W93–W99. <https://doi.org/10.2214/AJR.16.16638> (2017).
27. Chaibi, Y. *et al.* Fast 3D reconstruction of the lower limb using a parametric model and statistical inferences and clinical measurements calculation from biplanar X-rays. *Comput. Methods Biomech. Biomed. Engin.* **15**, 457–466. <https://doi.org/10.1080/10255842.2010.540758> (2012).
28. Tokunaga, K., Okamoto, M. & Watanabe, K. Implant orientation measurement after THA using the EOS X-ray image acquisition system. *Adv. Exp. Med. Biol.* **1093**, 335–343. https://doi.org/10.1007/978-981-13-1396-7_26 (2018).
29. Yamada, K. *et al.* Accuracies in measuring spinopelvic parameters in full-spine lateral standing radiograph. *Spine* **40**, E640–E646 (2015).
30. Vanhove, F. *et al.* Standardization of torsional CT measurements of the lower limbs with threshold values for corrective osteotomy. *Arch. Orthop. Trauma Surg.* **139**, 795–805. <https://doi.org/10.1007/s00402-019-03139-1> (2019).
31. Esposito, C. I. *et al.* Biplanar low-dose radiography is accurate for measuring combined anteversion after total hip arthroplasty. *HSS J.* **16**, 23–29. <https://doi.org/10.1007/s11420-018-09659-7> (2020).
32. Buck, F. M., Guggenberger, R., Koch, P. P. & Pfirrmann, C. W. A. Femoral and tibial torsion measurements with 3D models based on low-dose biplanar radiographs in comparison with standard CT measurements. *Am. J. Roentgenol.* **199**, W607–W612. <https://doi.org/10.2214/AJR.11.8295> (2012).
33. Ma, Z. *et al.* Assessing component orientation of total hip arthroplasty using the low-dose bi-planar radiographs. *BMC Musculoskelet. Disord.* **23**, 886. <https://doi.org/10.1186/s12891-022-05835-3> (2022).
34. Mayr, H. O. *et al.* Anteversion angle measurement in suspected torsional malalignment of the femur in 3-dimensional EOS vs computed tomography—a validation study. *J. Arthroplasty* **36**, 379–386. <https://doi.org/10.1016/j.arth.2020.07.058> (2021).
35. Girinon, F. *et al.* Quasi-automated reconstruction of the femur from bi-planar X-rays. *Comput. Methods Biomech. Biomed. Eng. Imaging Vis.* **8**, 529–537. <https://doi.org/10.1080/21681163.2020.1725915> (2020).
36. Ghelman, B., Kepler, C. K., Lyman, S. & DellaValle, A. G. CT outperforms radiography for determination of acetabular cup version after THA. *Clin. Orthopaed. Relat. Res.* **467**, 2362–2370 (2009).
37. De Pieri, E. *et al.* Subject-specific modeling of femoral torsion influences the prediction of hip loading during gait in asymptomatic adults. *Front. Bioeng. Biotechnol.* **9**, 1. <https://doi.org/10.3389/fbioe.2021.679360> (2021).

Acknowledgements

The authors would like to thank Antoine Poncet of the Clinical Research Centre of the Geneva University Hospitals for his help in designing the statistical analysis. This research was carried out with the support of EOS imaging that shared a research version of sterEOS allowing the extraction of anatomical landmark positions in addition to the standard geometric parameters of the pelvis and femur. EOS imaging did not influence the design of this study. This study was funded by the Division of Orthopaedic Surgery and Musculoskeletal Trauma Care of the Geneva University Hospitals (Geneva, Switzerland). The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.

Author contributions

Conceptualization: X.G, P.B, S.A, D.H; data curation: X.G, P.B, M.Z, S.B; formal analysis: X.G, S.A; investigation; methodology: X.G, P.B, S.A.; supervision: S.A, D.H; writing—original draft: X.G; writing—review and editing: P.B; P.D.G., M.Z, S.B., S.A., D.H.. All Authors have read and agreed to the published version of the manuscript.

Competing interests

The authors declare no competing interests.

Additional information

Supplementary Information The online version contains supplementary material available at <https://doi.org/10.1038/s41598-022-25997-3>.

Correspondence and requests for materials should be addressed to X.G.

Reprints and permissions information is available at www.nature.com/reprints.

Publisher's note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.



Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.

© The Author(s) 2022