Reproducibility of low-dose stereography measurements of femoral torsion after IM nailing of femoral shaft fractures and in intact femurs

J. Knafo *, T. Thelen, D. Verdier, L. Creppy, C. Tournier, T. Fabre

CHU Pellegrin, place Amélie-Raba-Léon, 33000 Bordeaux, France

ARTICLE INFO

Article history:
Received 27 August 2014
Accepted 21 March 2016

Keywords:
Femur fracture
Malunion
Femoral torsion
Low-dose stereography
EOS
Trauma

ABSTRACT

Introduction: Rotational malunion is a complication of intramedullary (IM) nailing for femur fractures. Symptoms can appear with 15° or more of axial deformity. None of the currently available measurement methods have a satisfactory reliability/irradiation ratio. The purpose of this study was to study the reproducibility of measuring femoral torsion with an EOS® low-dose stereography (LDX) system.

Hypothesis: LDX is a reproducible method for measuring post-traumatic femoral torsion.

Material and methods: The intra- and inter-observer reproducibility was studied in 45 patients who had a femoral fracture treated by IM nailing. Both the injured and contralateral healthy femurs were modelled. Bland–Altman plots were used to analyze the measurements made by three different observers (two orthopedic surgeons and one radiologist). For a given comparison, the interval between the upper limit of agreement (ULA) and lower limit of agreement (LLA) had to be within [−5°; 5°] for the examination to qualify as reproducible. Measurements were made by three observers (A, B, C) on the injured and healthy femur.

Results: With the fractured femurs (n = 39), the intra-observer [ULA; ULA] interval was [−16.295; 12.977]; it was [−18.475; 16.744] for the A–B pairing, [−13.316; 13.532] for the B–C pairing and [−17.839; 19.355] for the A–C pairing. With the healthy femurs (n = 37), the intra-observer [ULA; ULA] interval was [−7.909; 7.88]; it was [−11.924; 11.639] for the A–B pairing, [−12.654; 11.93] for the B–C pairing and [−11; 12.009] for the A–C pairing.

Discussion: The [ULA; ULA] intervals were greater than the [−5°; +5°] interval in all cases. LDX reproducibility is not sufficient for measuring femoral torsion after fracture or in healthy femurs. Observer experience, cohort size and the perfectible image quality are likely sources of bias. Conversely, the use of Bland–Altman plots and the multidisciplinary training of observers are major strengths of this study. Reproducibility will likely improve as the software is developed further and the image acquisition improves.

© 2016 Elsevier Masson SAS. All rights reserved.

1. Introduction

The gold standard treatment for femoral shaft fractures is intramedullary (IM) nailing. But the use of this technique increases the risk of rotational malunion.

The angle between the bicondylar line and the femoral neck axis, measured along the femur’s longitudinal axis (or femoral torsion) averages 15° internal rotation, with a physiological variation between 0° and 30°, and an average left-to-right asymmetry of 3.8° [0°–9.8°] [1]. Femoral malunion leading to horizontal plane deformity of 15° or more relative to normal (or relative to healthy contralateral leg) can lead to poor clinical outcomes [1–5].

Rotational malunion after femoral IM nailing is a common problem. More than 15° deformation has been reported in 19% to 28% of IM nailing cases [2,3]. Symptoms include discomfort during sports activities, hip or knee pain or stiffness, discomfort during walking (only for the most severe deformity) with insufficient compensation at the hip [6,7]. Progression to knee osteoarthritis has not been proven [3,7–9]. According to Braten, nearly 40% of patients had symptoms caused by the deformity and 14% had to be re-operated for this reason [2,10]. More recently, Karaman et al. [11] proposed that a greater than 10° deformity could explain why some patients have worse clinical outcomes.

For this reason, we need a tool that can be used to assess reliably of the situation before treatment that is low cost, exposes patients

* Corresponding author.
E-mail address: jeremie.knafo@chu-bordeaux.fr (J. Knafo).

http://dx.doi.org/10.1016/j.otsr.2016.03.020
1877-0568/© 2016 Elsevier Masson SAS. All rights reserved.
to the lowest possible dose of radiation and can be used in young subjects.

Several methods have been put forward to evaluate abnormal femoral torsion or rotational malunion. Multiple landmarks are used. They differ depending on the method used: line passing through center of femoral head and neck or anterior side of femoral neck proximally, bi-epicondylar line or posterior condylar line distally, horizontal plane or examination table plane or major axis of the femur as the reference plane.

Clinically, the method used to analyze hip rotation and/or palpate the greater trochanter is not reproducible [12]. The radiographic method described by Dunn and Ripstein requires an awkward patient position, several manual measurements and at the end, an indirect calculation of femoral torsion. There is a reliable ultrasonography method, but it is operator dependent and not widely used in France [1,2]. Computed tomography (CT) is the gold standard method, but there is a large variability in its reproducibility depending on the measurement method. Recent methods with image reconstruction and superimposition are reliable, but expose the patient to more radiation [12,13]. Magnetic resonance imaging also produces reliable measurements, but access is not universal, the cost is high and the measurement quality is unknown in the presence of metal fracture fixation devices (artefacts) [14].

A new radiological examination method called low-dose stereography (LDX) has been developed (EOS Imaging, Paris, France). This system generates reproducible three-dimensional reconstruction and measurements of spine and lower limb parameters that have been validated in the context of scoliosis surgery, and hip and knee osteoarthritis [15–20]. This system has several advantages: the radiation dose is up to 30 times less than a CT scan [18–21], images can be taken of standing patient, the cost is reasonable, and many anatomical parameters can be measured simultaneously [22]. However, LDX has not yet been studied in the context of trauma.

The goal of our study was to evaluate the reproducibility of this examination method to measure femoral torsion after IM nailing of the femoral shaft fractures. We hypothesized that reproducibility will be sufficient to use this imaging system for this indication in clinical practice.
2. Material and methods

Patients being treated in the orthopedics department of the Bordeaux University Hospital were enrolled prospectively. Patients were included if they were at least 18 years of age, had a unilateral femoral shaft fracture treated by IM nailing, and they could stand in the LDX enclosure. Patients were excluded if they had history of surgery or a prior femur fracture (on either side). Any exams in which the skeleton was not fully visible were excluded, because the goal was to assess the reproducibility of high-quality exams. The main reason for images to be excluded was incorrect patient positioning in the enclosure; this could be corrected before redoing the LDX examination without exposing the patient to damaging radiation levels. All patients received clear information about the reason for this examination.

Three observers prepared 3D models of the patients’ lower limbs using the SterEOS® software (EOS Imaging, Paris, France). After the LDX images were acquired, they were uploaded in the SterEOS® software. The observer located the center and diameter of the femoral head, axis of the femoral neck, center and diameter of the femoral condyles; the femur’s contour was adjusted manually at each step (Figs. 1 and 2).

The SterEOS® software calculated femoral torsion by measuring the angle between the femoral neck axis passing through the center of the femoral head and the line passing through the most posterior portion of the femoral condyles, along the femur’s longitudinal axis.

Observers A and B were orthopedic surgery residents and observer C was fellow in radiology, specializing in musculoskeletal imaging. The three observers were novices with the SterEOS® software.

For each examination, observer A carried out the 3D modelling twice at a one-month interval (series A1 and A2). Observers B and C each did the modelling once. Each examination included both lower limbs. Bland–Altman plots were used to evaluate the intra-observer reproducibility (series A1 and A2) and the inter-observer reproducibility (observers A, B, C) [23].

Within each series, the fractured and healthy femurs were analyzed separately. The lower and upper limits of agreement (LLA and ULA, respectively) and the 95% confidence intervals were recorded, along with the average of the differences between the measurements carried out.

![Fig. 2. Three-dimensional reconstruction generated by SterEOS®.](image)

This analysis was carried out with SAS software (version 9.3) by the Institut de santé publique, d’épidémiologie et de développement at the Bordeaux University. After consulting with statisticians, the number of subjects to enroll was set arbitrarily at 40 because of lack of data on this topic in published studies.

To evaluate the inter-observer reproducibility, pairs of observers were compared. Since observer A had done two series of measurements for each patient, we decided that if the intra-observer reproducibility was sufficient, one of observer A’s measurement series would be chosen randomly to evaluate inter-observer reproducibility. If the intra-observer reproducibility was not sufficient, the average of observer A’s observations would be used to evaluate inter-observer reproducibility.

When comparing the data series, the interval between the lower limit of agreement (LLA) and the upper limit of agreement (ULA) had to be within $[-5; 5]$ for the examination to qualify as reproducible. By definition, the [LLA; ULA] interval encompasses 95% of the differences measured between two series.

3. Results

Forty-five LDX exams were performed between November 1, 2011 and May 1, 2013. Seven patients were excluded secondarily because of a computer problem during image acquisition ($n = 2$), inability to stand in the enclosure ($n = 1$), or the femoral condyles could not be made out on the lateral views ($n = 4$). In all, 38 LDX exams were useable: 12 women and 26 men, average age of 34.7 years (range 18–82). One female patient had a bilateral fracture treated by IM nailing; her data was retained in the analysis of fractured femurs. This resulted in a cohort of 39 fractured femurs (18 right, 21 left) and 37 healthy femurs. Nine patients had undergone a second procedure to remove the IM nail before LDX was performed.

The Bland–Altman plots were used to generate average differences and ULA, LLA and 95% confidence intervals for comparing the two series. All the data are presented in Figs. 3 and 4. In both...
the intra-observer and inter-observer comparisons, the [LLA; ULA] intervals were above the [−5; 5] threshold.

4. Discussion

The goal of this study was to evaluate the agreement in the measurement of femoral torsion based on SterEOS® modelling of LDX examination of the lower body. For each set of comparisons, the [LLA; ULA] confidence limits were above the [−5; 5] interval (in degrees) that we had set as the clinically acceptable limit. Our hypothesis was not confirmed. The agreement is not sufficient and the reproducibility of the femoral torsion measurement was lacking in patients with a femoral shaft fracture treated by IM nailing. The agreement was also insufficient for normal femurs.

This can be explained by the sometimes average quality of the original images. Low-quality images were excluded as we could not identify anatomical landmarks on them. A meaningful number (n = 4) of images were excluded from the study for this reason (femoral condyles could not be located on the lateral view). Our objective was to measure the reproducibility of SterEOS® modelling based on high-quality EOS images, not to determine the overall efficiency of the process.

Even with images that were retained, it was sometimes difficult to identify the condyles or femoral neck because of radiographic superimposition on lateral views. Patient positioning seems to be perfectible in several cases. Chalbi et al. described how the femoral condyles are superimposed on lateral views and validated the “shifted feet” position (one foot ahead of the other) to offset these structures, making them more visible [24]. But we feel there is another important factor that comes into play when measuring femoral torsion: visualization of the proximal femur on the lateral view. Superimposition of the proximal femurs and pelvis, which occurred multiple times, made it difficult to accurately identify the contours of the femoral heads, femoral necks and metaphyses, even when the patient was in the shifted feet position. Improvements in the imaging acquisition still need to be made, especially for patient positioning.

The nature of the trauma, which led to significant deformity in some patients, may also have been a limitation, making the 3D modelling less reliable because the anatomy deviates from the “normal” anatomy that the system was designed to measure. However, the analysis of the measurements on healthy femurs also did not have sufficient agreement, which means that this technique is not reproducible based on the criteria set out in our study, even for non-deformed femurs. The observers’ limited experience in this domain and the number of patients enrolled may also have been limiting factors.

The strengths of this study are the use of three different observers who come from an orthopedic surgery and musculoskeletal radiology background. Use of Bland–Altman plots was also a strength for this study. We chose this method because it is superior to the correlation tests typically used in this type of study. Correlation tests measure the amount of dependence between several series of measurement, while the Bland–Altman method also measures the agreement between the values in these series. This means that the proximity of the values for each pair of measurements was measured, not only their dependence [23]. Only a few studies have used this statistical method; those that concluded positively about LDX reproducibility in other indications did not. This may also be the reason why our results are not consistent with previous studies.

Another factor is improvements in the SterEOS® software, which was designed to carry out modelling and measurements based on stereography images. This software has undergone additional developments since our study was performed. The modelling steps have been simplified and an algorithm to help locate the femoral condyles was included in the software. This algorithm aims to help non-experienced users model the femurs without reversing the medial and lateral condyles [25].

Based on our findings, we do not recommend using LDX in daily practice to measure femoral torsion, either in the trauma context after fracture or in orthopedics. In fact, the lack of reproducibility of this process between our different observers drove us to continue in use CT as the gold standard examination when a reliable measurement of femoral torsion is needed. Improved imaging acquisition protocols, along with a new version of the modelling software, may make it possible for LDX to be used reliably for this type of measurement in the future.

Disclosure of interest

The authors declare that they have no competing interest.

References

malalignment

Orthotème

versus

comparison

normal
ture:

ubn.ru.nl/handle/2066/19538

pen

Orthop

arthritis.

Than

Dubousset

Jaarsma

Karaman

Jaarsma

O,

Guggenberger

T.

Serrurier

et

Am

Roentgenol


Folinais

D,

Thelen

P,

Delin

C,

et

al.


