 Acta Chir Orthop Traumatol Cech. 89, 2022, No. 1
p. 75–80

ORIGINAL PAPER
PŮVODNÍ PRÁCE

Reliability of Evaluation of the Surgeon-Dependent Factors Affecting Mechanical Failure after Intertrochanteric Femur Fracture Treatment

Spolehlivost hodnocení na operatérovi závislých faktorů, jež ovlivňují mechanické selhání v léčbě intertrochanterické zlomeniny femuru

A. TURGUT, M. FILIBELI, M. KUMBARACI, E. BILGIN, S. HANCIOĞLU, Ö. KALENDERER

Tepecik Training and Research Hospital, Department of Orthopaedics and Traumatology, İzmir, Turkey

ABSTRACT

PURPOSE OF THE STUDY

The intertrochanteric femur fractures seen in the elderly population are usually treated surgically. Mechanical failure of this treatment is a serious complication and to prevent this, many assessment factors have been described such as tip-apex distance, reduction quality etc. The aim of this study is to evaluate the intra- and inter-observer reliability of modifiable factors evaluated after the treatment of intertrochanteric fractures.

MATERIAL AND METHODS

The early post-operative radiographs of sixty-two patients treated with PFN-A were evaluated. Six observers with 8 to 27 years of experience in hip trauma surgery were asked to measure tip apex distance (TAD), calcar referenced TAD (calTAD), collo-diaphyseal angle (CDA) and evaluate the position of helical blade and also fracture reductions according to Baumgaertner and Chang criteria on these radiographs. Fleiss kappa and intra-class correlation coefficient (ICC) values were calculated for inter-observer and intra-observer reliability assessment, respectively.

RESULTS

Inter-observer reliability values for both evaluations were moderate (fleiss κ: 0.417–0.455) for TAD measurements, moderate (fleiss κ: 0.418–0.458) for calTAD measurements, fair (fleiss κ: 0.302–0.288) for CDA measurements, substantial (fleiss κ: 0.606–0.631) for antero-posterior zone evaluation, moderate (fleiss κ: 0.550–0.546) for lateral zone assessment, fair-moderate (fleiss κ: 0.353–0.453) for Baumgaertner reduction quality assessment, and fair-moderate (fleiss κ: 0.365–0.456) for Chang reduction quality assessment. Intra-observer reliabilities were found to be moderate to good for TAD and calTAD measurements, moderate for CDA measurements, good to excellent for AP and lateral zone and Baumgaertner reduction quality criteria assessments and good for Chang reduction quality criteria assessment.

DISCUSSION

There is no large data on the subject of the present study in the literature. There is a single article evaluating the reliability of more than one evaluation criteria in which only two observers' findings were evaluated in the literature. As in our study, more accurate inferences can be made with the increase in the number of observers. Unlike the literature, the inter-observer reliabilities of TAD and calTAD which have a very important place in the postoperative evaluation of these fractures were found as "moderate" instead of "excellent" in our study,

CONCLUSIONS

It was determined that only the inter-observer reliability of antero-posterior zone evaluation was substantial. The inter-observer reliability of all other measurements and evaluations were fair to moderate.

Key words: Intertrochanteric fractures, reliability, measurements, reduction quality; TAD, calTAD.

INTRODUCTION

Intertrochanteric femur fractures are commonly seen in elderly osteoporotic patients (6, 13). Due to the high healing potential, these fractures are generally treated with fixation. In the fixation of these fractures, extramedullary methods such as dynamic hip screw and intramedullary methods such as proximal femoral nail anti-rotation (PFN-A) (Synthes GmbH, Oberdorf, Switzerland) are frequently used (12). Mechanical failure after treatment is a serious cause of morbidity and possible mortality (15).

One of the most important mechanical complications encountered after fixation of intertrochanteric fractures is the cut-out, which can be defined as the collapse of the neck-shaft angle into varus, leading to extrusion, or so-called cut-out, of the cephalic implant from the femoral head (2). Factors that may be responsible for the cut-out complication are divided into two groups as modifiable and non-modifiable. Modifiable factors are; the quality of the reduction achieved (1, 3), the appropriateness of the tip apex distance (TAD) (2) and the quadrant where the cephalic implant is placed (4). The methods used in the treatment of trochanteric fractures

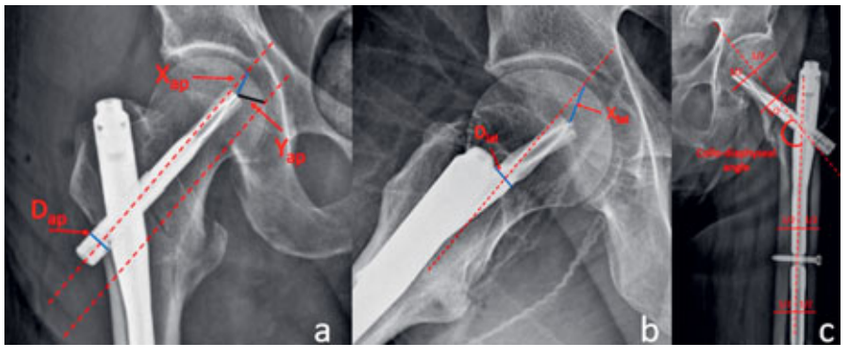

Fig. 1. Measurement method of the tip apex distance and calcar referenced tip apex distance on AP radiograph (1a), tip apex distance measurement method on lateral radiograph (1b), measurement method of collo-diaphyseal angle (1c)
TAD= [Xap x (D true/Dap)] + [Xlat x (Dtrue/Dlat)]
calTAD= [Yap x (D true/Dap)] + [Xlat x (Dtrue/Dlat)]
(D true: known true diameter of the helical blade).

Table 1. Baumgaertner reduction quality criteria

I- Alignment
- a. Antero-posterior view: normal or slight valgus neck-shaft angle
- b. Lateral view: less than 20° of angulation

II- Displacement
- a. Antero-posterior view: less than 4 mm of displacement of any fragments
- b. Lateral view: less than 4 mm of displacement of any fragments

***Reduction quality criteria**
Good: both criteria met
Acceptable: only one criterion met
Poor: neither criterion met

Table 2. Chang reduction quality criteria

	Score
I- Alignment	
a. Antero-posterior view: normal or slight valgus neck-shaft angle	1
b. Lateral view: less than 20° of angulation	1
II- Displacement	
a. Antero-posterior view: neutral or positive medial cortical support	1
b. Lateral view: smooth anterior cortical contact	1
***Reduction quality**	
Excellent	4
Acceptable	2–3
Poor	0–1

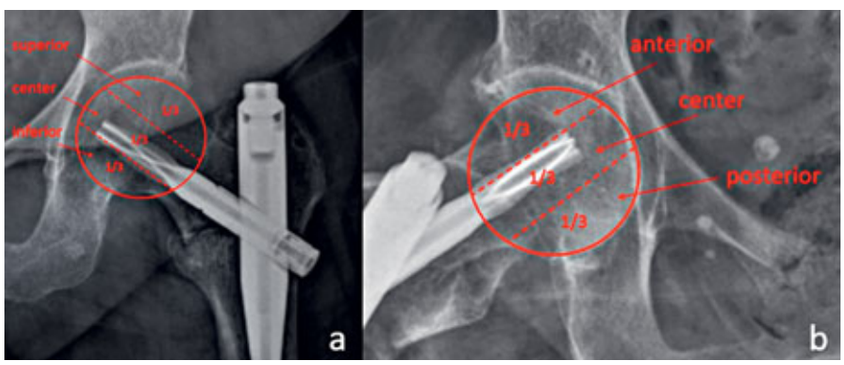

Fig. 2. Determination of implant zones on antero-posterior (2a), and lateral radiographs (2b).

and complications in osteoporotic elderly patients having one or more comorbidities are of great interest to many authors. The main parameters used in the statistical analysis of these studies and providing conclusions are the evaluations of the modifiable factors that are responsible for mechanical failure. Therefore, it is very important that the evaluation of these parameters have high intra- and inter-observer reliability. In the literature, there is very limited data reported on the reliability of modifiable parameters evaluated after intertrochanteric fracture treatment. Most of these evaluations were inferred from the results of two observers.

The aim of this study is to evaluate the intra- and inter-observer reliability of modifiable factors evaluated after treatment of intertrochanteric fractures.

MATERIAL AND METHODS

After obtaining the local ethical committee approval (number:2020/11-12) the information of patients with a hip fracture between June 2018 and January 2019 was obtained from the institutional archive. The inclusion criteria was; intertrochanteric hip fractures which were treated with PFN-A with an antero-posterior/lateral hip radiographs. The patients who were treated with arthroplasty were excluded. There were totally 62 patients who were included in the study. The early postoperative AP and lateral hip radiographs of these patients were saved in a Digital Imaging and Communications in Medicine (DICOM) format. After providing detailed information about the evaluation of modifiable factors to six observers with 8 to 27 years of experience in hip trauma surgery, they were asked to measure TAD, calcar referenced TAD (calTAD) (9), collo-diaphyseal angle (CDA) using the RadiANT DICOM Viewer 2.2.9® (Medixant, Poznan, Poland) program. Observers were asked to evaluate the position of helical blade and also fracture reductions according to Baumgaertner (1) and Chang criteria (3) on these radiographs. Four weeks after the completion of the first evaluation, the observers were asked to re-evaluate the same parameters on reordered radiographs of the same patient group. The observers should have completed their evaluations in two weeks. The measurement methods of TAD, CalTAD and CDA

Table 3. Inter-observer reliabilities among the first and second evaluations

	First evaluation Fleiss kappa (95% confidence interval)	Second evaluation Fleiss kappa (95% confidence interval)
TAD	0.417 (0.353–0.481)	0.455 (0.390–0.519)
CalTAD	0.418 (0.354–0.482)	0.458 (0.393–0.522)
CDA	0.302 (0.237–0.366)	0.288 (0.224–0.352)
AP zone	0.606 (0.548–0.663)	0.631 (0.575–0.688)
Lateral zone	0.550 (0.500–0.601)	0.546 (0.496–0.595)
Baumgaertner criteria	0.353 (0.303–0.403)	0.453 (0.405–0.502)
Chang criteria	0.365 (0.314–0.416)	0.456 (0.407–0.505)

TAD: tip apex distance, calTAD: calcar referenced tip apex distance, CDA: collo-diaphyseal angle, AP: antero-posterior

are shown in Figure 1a–c. The position of the helical blade was evaluated as, superior-center-inferior on AP radiographs and anterior-center-posterior on lateral radiographs (Fig. 2a, b). The reductions were grouped as, good-acceptable-poor according to the Baumgaertner criteria (1) (Table 1) and excellent-acceptable-poor according to the Chang criteria (3) (Table 2). The measurements of TAD and calTAD was grouped as ≤25 and >25 millimeters (mm) and CDA measurements were grouped as ≤127° and >127° (varus-valgus, respectively) for statistical analysis.

TAD and calTAD distances measured in the first and second evaluations were compared with respect to AP zones and between each-other.

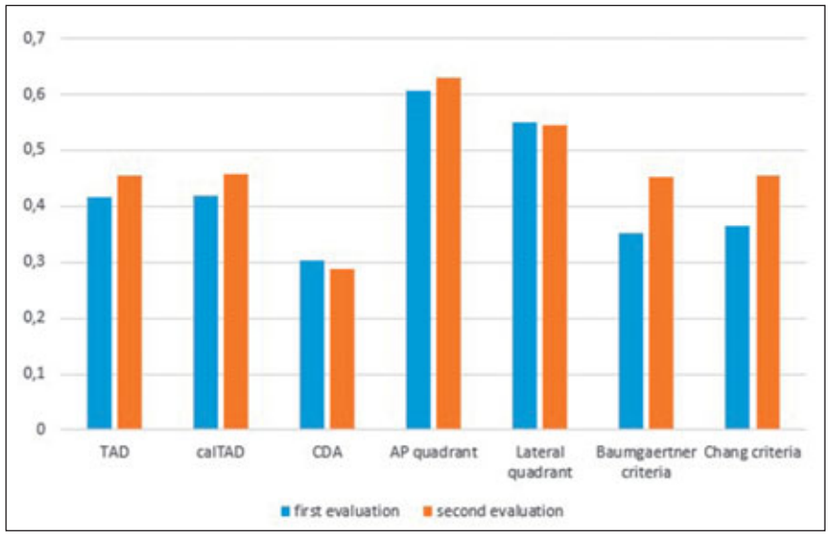

Fig. 3. Bar chart showing inter-observer reliability values of each variable.

Statistical analyses were carried out using the IBM Statistical Package for Social Sciences (SPSS) for Windows, version 24.0 (IBM Corp., Armonk, New York, USA). Intra-class Correlation Coefficient (ICC) values were calculated for the evaluation of intra-observer reliability. Fleiss κ values other than ICC were calculated for the evaluation of inter-observer reliability by using Microsoft® Office Excel, because there were more than two observers. As suggested by Altmann, the κ values were interpreted as follows: below 0.20 as slight agreement, 0.21–0.40 as fair, 0.41–0.60 as moderate, 0.61–0.80 as substantial, and over 0.80 as very good agreement (10). An ICC value less than 0.5, between 0.5 and 0.75, between 0.75 and 0.9, and greater than 0.90 were accepted as poor, moderate, good, and excellent reliability, respectively (8). Shapiro wilk test was used for evaluation of the normality of continuous data. If the data was distributed normally T-test or ANOVA test was used for comparison according to the group number. If the data was non-normally distributed, Mann-Whitney U test or Kruskal-Wallis test was used for comparison according to the group number. A p value 85 below than 0.05 was accepted as statistically significant.

RESULTS

Forty-two (67.7%) of the patients were female. The mean age was 71.6±8.8 (58–86) years. Interobserver reliability values for both measurements were moderate (fleiss κ: 0.417–0.455) for TAD measurements, moderate (fleiss κ: 0.418–0.458) for calTAD measurements, fair (fleiss κ: 0.302–0.288) for CDA measurements, substantial (fleiss κ: 0.606–0.631) for AP zone evaluation, moderate (fleiss κ: 0.550–0.546) for lateral zone assessment, fair-moderate (fleiss κ: 0.353–0.453) for Baumgaertner reduction quality assessment, and fair-moderate (fleiss κ: 0.365–0.456) for Chang reduction quality assessment (Table 3, Fig. 3).

Intra-observer reliabilities were found to be moderate good for TAD and calTAD measurements, moderate for CDA measurements, good to excellent for AP and lateral zone and Baumgaertner reduction quality criteria assessments and good for Chang reduction quality criteria assessment (Table 4, Fig. 4).

When TAD and calTAD measurement values in AP zones were evaluated, calTAD measurement values in superior and middle zones were found to be higher than

Table 4. Intra-observer reliabilities of each factor

	TAD ICC (95% confidence interval)	CalTAD ICC (95% confidence interval)	CDA ICC (95% confidence interval)	AP zone ICC (95% confidence interval)	Lateral zone ICC (95% confidence interval)	Baumgaertner criteria ICC (95% confidence interval)	Chang criteria ICC (95% confidence interval)
Observer 1	0.808 (0.682–0.884)	0.803 (674–0.882)	0.608 (0.349–0.764)	0.953 (0.922–0.972)	0.946 (0.910–0.967)	0.823 (0.707–0.894)	0.805 (0.696–0.878)
Observer 2	0.850 (0.751–0.910)	0.670 (0.452–0.801)	0.596 (0.330–0.757)	0.933 (0.889–0.960)	0.929 (0.881–0.957)	0.807 (0.699–0.879)	0.808 (0.681–0.884)
Observer 3	0.876 (0.793–0.925)	0.802 (0.671–0.880)	0.586 (0.313–0.751)	0.820 (0.701–0.891)	0.818 (0.698–0.890)	0.871 (0.786–0.922)	0.811 (0.686–0.886)
Observer 4	0.785 (0.644–0.871)	0.804 (0.675–0.882)	0.568 (0.283–0.740)	0.959 (0.932–0.975)	0.955 (0.926–0.973)	0.916 (0.860–0.949)	0.896 (0.827–0.937)
Observer 5	0.771 (0.620–0.862)	0.770 (0.619–0.862)	0.521 (0.205–0.711)	0.911 (0.852–0.946)	0.812 (0.688–0.887)	0.868 (0.780–0.920)	0.834 (0.724–0.900)
Observer 6	0.750 (0.584–0.849)	0.621 (0.370–0.771)	0.555 (0.262–0.732)	0.813 (0.689–0.887)	0.808 (0.681–0.884)	0.808 (0.682–0.884)	0.817 (0.697–0.890)

TAD: tip apex distance, calTAD: calcar referenced tip apex distance, CDA: collo-diaphyseal angle, AP: antero-posterior

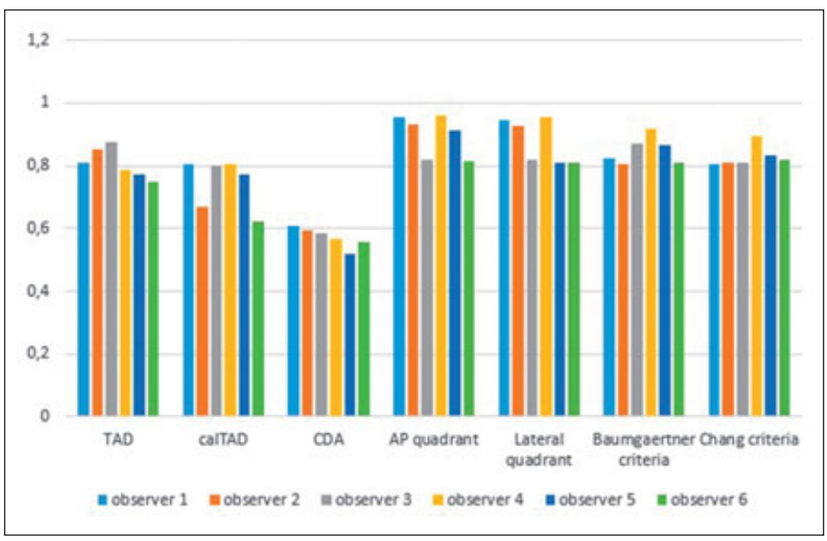

Fig. 4. Bar chart showing intra-observer reliability values among six observers.

TAD measurement values. In the inferior zone, calTAD measurement values were found to be lower than TAD measurement values. CalTAD values were statistically significantly lower in the inferior zone than other zones ($p<0.001$). When comparisons were made for each zone, a statistically significant difference was found in the second evaluations, even if not in the first evaluations (p values for the superior, center and inferior zones were 0.04, <0.001 and 0.01, respectively). The lowest TAD measurement values were detected in the center zones (p values are 0.079 and 0.004 for the first and second evaluation, respectively) (table 5).

DISCUSSION

In making conclusions about the results of all studies on internal fixation of intertrochanteric fractures, the most important factors to be considered are the quality of the reduction, the values of TAD measurements and the location of the cephalic implant in the femoral head, regardless of the implant used. As they are surgeon dependent modifiable factors, it is important to investigate them. If these factors are applied as recommended, the prognosis of the fracture will be better, and the morbidity and mortality rates related to the fracture can be reduced. The purpose of this present study was to investigate the reliability of the evaluations of these important factors. As a result of this study, it was determined that the evaluation of the implant's placement zone in the head was at the highest levels of both intra- and inter-observer reliability, and the lowest reliability was seen in the decision making between varus or valgus fracture position. Inter-observer reliabilities of Baumgaertner and Chang criteria used in the evaluation of re-

Table 5. Comparison of the TAD and calTAD values according to the coronal plane quadrants

	First evaluation			Second evaluation		
	TAD mean±std dev. (min–max)	calTAD mean±std dev. (min–max)	p value	TAD mean±std dev. (min–max)	calTAD mean±std dev. (min–max)	p value
Superior zone	28.5±10.3 (12.5–49.4)	33.6±10.8 (21.1–52.1)	0.20**	29.8±9.2 (10.3–45.8)	35.9±8.4 (19.3–51.1)	**0.04****
Center zone	26.1±6.7 (10.3–51.6)	29.2±6.7 (13.7–54.1)	**<0.001****	26.1±6.9 (9.9–56.6)	28.9±6.5 (15.1–59.5)	**<0.001****
Inferior zone	27.8±6.7 (12.9–46.7)	26.3±5.9 (7.8–45.5)	0.165**	27.9±6.6 (11.9–44.6)	25.7±5.4 (9.4–40.7)	**0.01****
p value	0.079*	**<0.001***		**0.004***	**<0.001***	

*comparison of the values in superior, center and inferior zones, **comparison of the values of TAD and calTAD in the same zone
TAD: tip apex distance, calTAD: calcar referenced tip apex distance, std dev: standard deviation, min: minimum, max: maximum

duction quality were found to be fair to moderate and intra-observer reliabilities were found to be good to excellent.

The first study on reliability of predictors for screw cut-out in intertrochanteric fractures was published in 2012 by De Bruijin et al. (5). They investigated the reliability of the TAD measurement, the Cleveland femoral head dividing system, the three-grade classification system of Baumgaertner for fracture reduction, and the Arbeitsgemeinschaft für Osteosynthesefragen (AO) classification system. The authors reported 0.81 and 0.89 for TAD (single measure and average measure values, respectively), 0.46 for screw position and 0.48 for fracture reduction (ICC or Weighted K). These results contradict our findings except than reduction assessment. As a result of our study, it was found that the reliability of determining the location of the implant in the head was the highest (moderate to substantial), whereas the inter-observer reliability of the TAD and calTAD measurements were moderate and fair, respectively. Another important difference between the two studies was that there were two observers in De Bruijin et al.'s study and six observers in the present study. As the number of observers increase, it can be predicted that the parameter whose reliability is tested will be evaluated more accurately.

A TAD value above or below 25 mm is a parameter that is frequently handled in the literature (14). Baumgaertner et al. stated that they evaluated the intra- and inter-observer reliability of the measurement method as a study team (four observers) in their study in which they first described TAD and concluded that this method was reproducible. They did not specify any κ or ICC values (1). As a result of the measurements made by two observers, Kashigar et al. reported the ICC values for TAD and calTAD measurements as 0.915 and 0.901, respectively (7). In the results of present study, the inter-observer reliabilities of the TAD and calTAD measurements were found to be fleiss κ: 0.418–0.458 and fleiss κ: 0.302–0.288, respectively. The reason for this lower reliability may be the need of millimetric measurement in TAD method and the evaluation of additional reference points such as the center of the helical blade tip and the center of the femoral head. With these results, it is not possible to indicate that the reliability of the TAD measurement is high enough.

Evaluation of the reduction quality is also very important in predicting the prognosis of intertrochanteric fractures. Since Chang reduction evaluation criteria were established relatively recently, there is only one study, performed by Mao et al. (11), in the literature comparing these criteria and Baumgaertner criteria. Based on only two observers' evaluations, the authors reported the reliability for TAD, blade position, Baumgaertner reduction criteria and Chang reduction criteria as 0.888 (ICC), 0.420 (κ), 0.589 (κ) and 0.731 (κ), respectively. In the study, the authors stated that the reason for the high reliability of Chang reduction criteria were use of a four-point system, the importance of positive medial cortical support and the necessity of 4 mm displacement measurement in Baumgaertner reduction criteria. Similarly, in the present study, the reliability of the Chang reduction criteria was higher than the evaluation of the Baumgaertner reduction criteria (0.365 vs. 0.353 for the first evaluation, 0.456 vs. 0.453 for the second evaluation, respectively).

The calTAD measurement method was developed since the inferior location of the cephalic implant in the femoral head in the coronal plane is predicted as a factor that increases the stability (9). As can be seen from our study results, calTAD measurement values are quite high if the cephalic implant is determined as inserted in the superior and center zones in the coronal plane compared to TAD measurements, and lower in the inferior zone. When we think with a reverse logic, classical TAD measurements result in above the recommended values even if the subchondral placement is applied to provide the necessary stability for inferiorly positioned implants. If we apply the calTAD measurement method for each coronal zone location, high measurement results will be encountered for implants with the recommended subchondral distance in superior and center zones. For all the reasons, we think that calTAD measurement provide more accurate evaluations in inferior placement of the cephalic implants. In contrary, accurate values in superior and central placements can be obtained with the help of traditional TAD measurements.

The most important limitation of this study is that intertrochanteric fractures were not followed up and the parameters evaluated in patients with mechanical failure were not compared with those who did not. For the first time in the literature, evaluation with six observers made up the most important strength of this current study.

CONCLUSIONS

As a conclusion, it was determined that only the inter-observer reliability of antero-posterior zone evaluation was substantial. The inter-observer reliability of all other measurement and evaluations were fair to moderate. According to the findings of present study we can state that; in studies where the relationship of factors with mechanical failure is evaluated, it will be appropriate to make the measurements carefully and if possible, by more than two people and to include these measurements and evaluations in the analysis by considering the average.

References

1. Baumgaertner MR, Curtin SL, Lindskog DM. Intramedullary versus extramedullary fixation for the treatment of intertrochanteric hip fractures. Clin Orthop Relat Res. 1998;348:87–94.
2. Baumgaertner MR, Curtin SL, Lindskog DM, Keggi JM. The value of the tip-apex distance in predicting failure of fixation of peritrochanteric fractures of the hip. J Bone Joint Surg Am. 1995;77:1058–1064.
3. Chang SM, Zhang YQ, Ma Z, Li Q, Dargel J, Eysel P. Fracture reduction with positive medial cortical support: a key element in stability reconstruction for the unstable pertrochanteric hip fractures. Arch Orthop Trauma Surg. 2015;135:811–818.

4. Cleveland M, Bosworth DM, Thompson FR, Wilson HJ, Ishizuka T. A ten-year analysis of intertrochanteric fractures of the femur. J Bone Joint Surg Am. 1959;41:1399–1408.
5. De Bruijn K, den Hartog D, Tuinebreijer W, Roukema G. Reliability of predictors for screw cutout in intertrochanteric hip fractures. J Bone Joint Surg Am. 2012;94:1266–1272.
6. FeldinszkÁ J, Jacko P, Barinka J, Kilian M, Šimko P. Comparison of 2-screw implant and antirotational blade implant in treatment of trochanteric fractures. Acta Chir Orthop Traumatol Cech. 2020;87:268–272.
7. Kashigar A, Vincent A, Gunton MJ, Backstein D, Safir O, Kuzyk PR. Predictors of failure for cephalomedullary nailing of proximal femoral fractures. Bone Joint J. 2014;96-B:1029–1034.
8. Koo TK, Li MY. A guideline of selecting and reporting ıntraclass correlation coefficients for reliability research. J Chiropr Med. 2016;15:155–163.
9. Kuzyk PR, Zdero R, Shah S, Olsen M, Waddell JP, Schemitsch EH. Femoral head lag screw position for cephalomedullary nails: a biomechanical analysis. J Orthop Trauma. 2012;26:414–421.
10. Landis JR, Koch GG. The measurement of observer agreement for categorical data. Biometrics. 1977;33:159–174.
11. Mao W, Ni H, Li L, He Y, Chen X, Tang H, Dong Y. Comparison of Baumgaertner and Chang reduction quality criteria for the assessment of trochanteric fractures. Bone Joint Res. 2019;8:502–508.
12. Sambandam SN, Chandrasekharan J, Mounasamy V, Mauffrey C. Intertrochanteric fractures: a review of fixation methods. Eur J Orthop Surg Traumatol. 2016;26:339–353.
13. Turgut A, Kalenderer Ö, Günaydın B, Önvural B, Karapınar L, Ağuş H. Fixation of intertrochanteric femur fractures using Proximal Femoral Nail Antirotation (PFNA) in the lateral decubitus position without a traction table. Acta Orthop Traumatol Turc. 2014;48:513–520.
14. Turgut A, Kalenderer Ö, Karapınar L, Kumbaracı M, Akkan HA, Ağuş H. Which factor is most important for occurrence of cutout complications in patients treated with proximal femoral nail anti-rotation? Retrospective analysis of 298 patients. Arch Orthop Trauma Surg. 2016;136:623–630.
15. Yam M, Chawla A, Kwek E. Rewriting the tip apex distance for the proximal femoral nail anti-rotation. Injury. 2017;48:1843–1847.

Corresponding author:
Ali Turgut, Assoc. Prof,
Tepecik Training and Research Hospital
Department of Orthopaedics and Traumatology
İzmir, Turkey
E-mail: draliturgutort@yahoo.com.tr