CrossMark

ORIGINAL PAPER

What is the most reproducible classification system to assess tibial plateau fractures?

Angélica Millán-Billi[1] · Mireia Gómez-Masdeu[1] · Eduard Ramírez-Bermejo[1] ·
Maximiliano Ibañez[1] · Pablo Eduardo Gelber[1,2]

Abstract

Purpose Traditional classification systems for tibial plateau fractures (TPF) are based on simple radiographs, and intra- and inter-observer variability is low. The aim was to assess intra- and inter-observer variability using traditional systems and some recently described classification systems of TPF in the interpretation of standard radiographs and bidimensional (2D) and tridimensional (3D) computed tomography (CT).

Methods We studied all patients at two centres who underwent TPF surgery over a three-year period. Demographic data (age, sex, BMI) and mechanism of injury were recorded. Four observers classified each TPF according to the Schatzker, AO, Luo, modified Duparc and Khan classification systems. We calculated intra- and inter-observer variability using the Kappa test.

Results A total of 112 (71 males) patients were included. Mean age was 47.1 years (range 21–86) and mean BMI was 25.2 ± 3.6. Intra- and inter-observer variability was 0.95 and 0.62 for AO, 0.87 and 0.65 for Schaztker, 0.86 and 0.73 for Luo, 0.56 and 0.37 for the modified Duparc, and 0.43 and 0.25 for Khan classifications.

Conclusions Although previous training could be needed, AO, Schatzker and Luo classifications showed a good reproducibility of TPF assessment from a combination of standard radiographs and 2D and 3D CT images. The results using the Modified Duparc and Khan classifications were less favourable and their use is not therefore recommended.

Keywords Tibial plateau · Fracture classification · Interobserver

Introduction

The tibial plateau is one of the most critical load-bearing areas in the human body. Fractures of the plateau affect knee alignment, stability, and motion. Early detection and appropriate treatment of these fractures are critical to minimize patient disability and reduce the risk of documented complications, particularly posttraumatic arthritis [1]. As the anatomy of the tibial plateau is complex, careful surgical planning with standard radiographies and computed tomography (CT) is essential. In general, surgeons make treatment decisions by assessing the radiology and grading the fracture pattern based on several classification systems [2].

A good classification system should be simple but accurate, show good reproducibility, and be based on clinically relevant data. Several systems have been described to classify tibial plateau fractures (TPF) [3–9]. The most widely used are the Schatzker [3] and AO [10] systems but concern has been raised over the accuracy and reproducibility of these two classifications [8, 9, 11]. Several new classification systems have since been developed, but few of them have been compared to date [12–14].

The purpose of this study was to compare the five most commonly used classification systems for TPF by evaluating their intra- and inter-observer reproducibility with standard radiography and CT scan. The hypothesis was that all five systems would have a low degree of inter-observer agreement with acceptable intra-observer variability.

✉ Pablo Eduardo Gelber
  personal@drgelber.com

1 Hospital de la Santa Creu i Sant Pau, Universitat Autònoma de Barcelona, Sant Quintí 89, 08041 Barcelona, Spain

2 ICATME-Hospital Universitari Dexeus, Universitat Autònoma de Barcelona, Sabino de Arana 5-19, 08028 Barcelona, Spain

Springer

Material and methods

We retrospectively studied all patients who underwent TPF surgery at two centres between 2013 and 2015. We recorded demographic data (age, gender and body mass index (BMI)) as well as the injury mechanism and knee side.

We only included cases with available anteroposterior and standard lateral radiographs and bidimensional (2D) and tridimensional (3D) CT scan reconstructions. We excluded patients whose radiographs were not performed in the index hospitals and patients with isolated spine avulsion fractures.

Four observers (one senior orthopaedic surgeon, one junior orthopaedic surgeon, one orthopaedic surgeon fellow and one senior resident) evaluated the radiological findings. The observers analyzed and classified the fracture using five different systems: AO [15], Schatzker [3], Luo [2], Khan [9] and the revised Duparc [16] systems. A diagrammatic scheme with a written description of these four classifications was also provided. All the observers evaluated the anteroposterior and lateral views of the standard radiography, and the two most representative views of axial, sagittal and coronal images of the CT and their 3D reconstruction. An independent observer had previously selected each of the CT views. The observers evaluated the images on two occasions, eight weeks apart, to determine intraobserver reliability. Their second evaluation was used to assess the interobserver agreement. None of the observers had been involved in the treatment of these patients' fractures.

Informed consent was obtained from all individual participants included in the study.
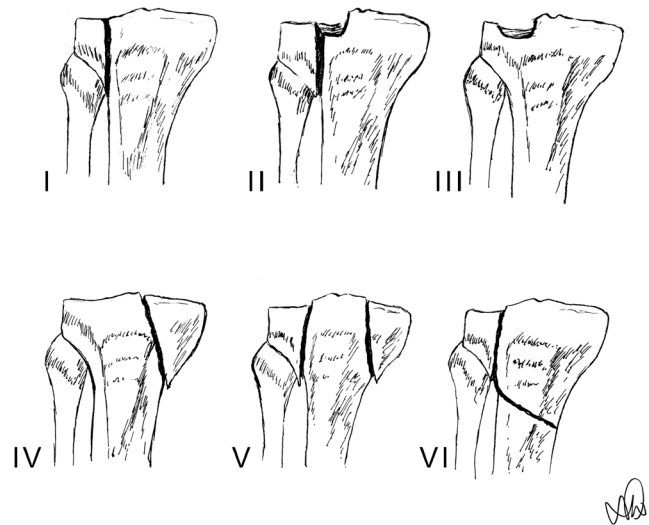
Classifications systems

In the AO classification system, the TPF corresponds to number 41 [15]. This system classified the fractures into **A** (extra articular), **B** (partial articular), and **C** (complete articular), with subtypes in each group (Fig. 1).
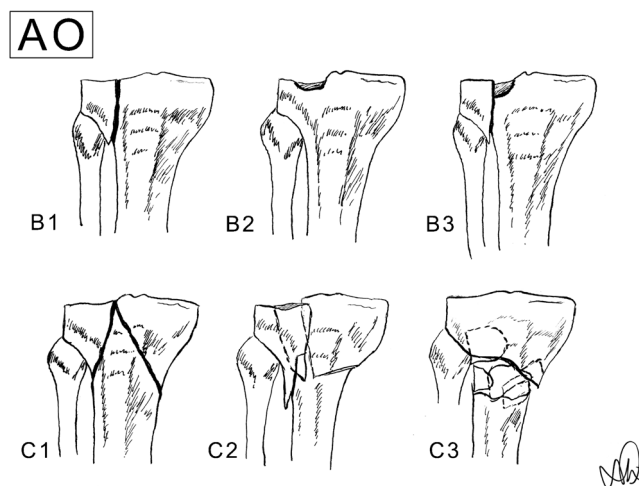
Fig. 2 Schaztker classification of TPF

The Schaztker classification is based on severity of the fracture [3], classifying TPF into six types: **I** - wedge-shaped pure cleavage fracture of the lateral tibial plateau, **II** - split and depression of the lateral tibial plateau, **III -** pure depression of the lateral tibial plateau, **IV** - pure depression of the medial tibial plateau, **V** - involving both tibial plateau regions, and **VI** - fracture through the metadiaphysis of the tibia (Fig. 2). The three-Column Luo's classification system is based in CT and 3D reconstructions, selecting views containing most parts of the fibular head and condylar spine on the axial CT views of the tibial plateau and dividing it into three columns (medial, lateral and posterior) (Fig. 3) [2]. The classification described by Khan et al. [9] grouped the fractures into lateral, medial, posterior, anterior, rim, bicondylar and subcondylar. In each group, the fracture was subclassified with numbers depending on the characteristics of the fracture, creating an alphanumeric system (Fig. 4).

Fig. 1 AO classification system of TPF
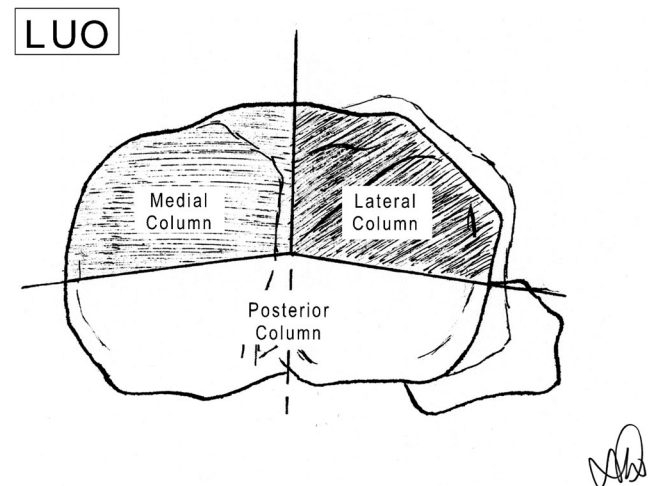
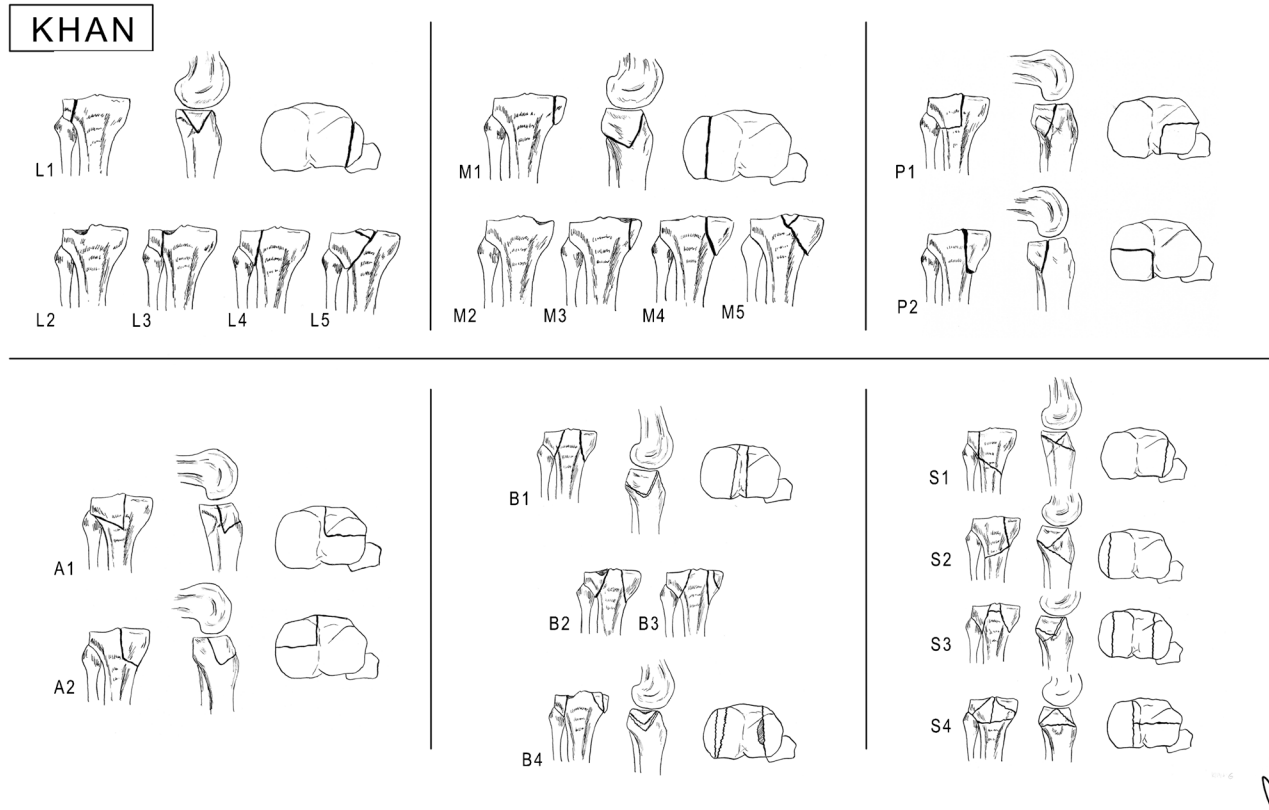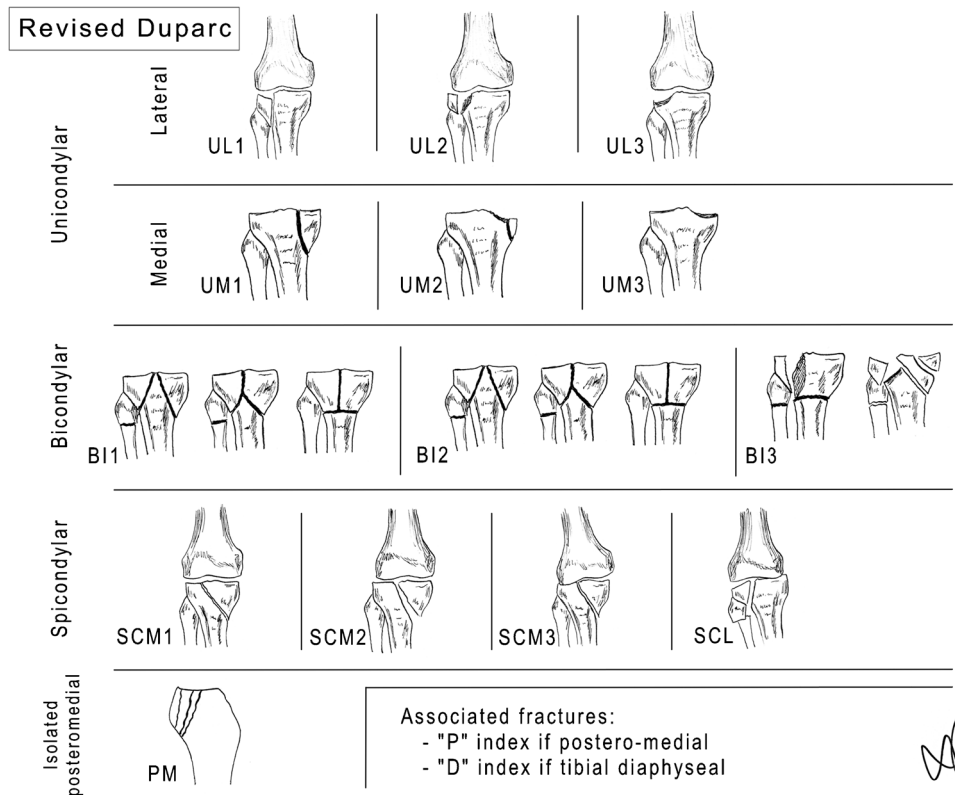Fig. 3 Drawing of the three-Column Luo's classification system of TPF

Fig. 4 Khan classification of TPF

Fig. 5 Revised Duparc classification of TPF

Table 1 Reproducibility of the five classification systems

Classification	A - Interobserver (k)	B - Intraobserver (k)
AO	0.62	0.85
Schatzker	0.65	0.87
Luo	0.73	0.86
Revised Duparc	0.37	0.56
Khan	0.25	0.43

The revised Duparc classification, proposed by Gicquel et al. [16], is based on groups of unicondylar, bicondylar, spinocondylar and isolated posteromedial fractures with acronyms for each fracture (Fig. 5).

Statistical analysis

Statistical analysis was performed using SPSS 19 (SPSS, Chicago, IL). Categorical variables are expressed as percentages and frequencies. Means and standard deviations as well as median, minimums, and maximums were calculated for each continuous variable.

The kappa coefficient (K) [17] was calculated to analyze the reliability classification system made by the same observer on separate occasions (intra-observer reliability) or by five different observers (inter-observer reliability).

The K statistic reflects how many responses the observers agreed on and how many agreements occurred by chance [18]. A 100% agreement had a value of 1.00 (maximum), while agreement attributed to chance had a value of 0 (minimum). The values were interpreted according to Landis and Koch [19]: <0.21 slight, 0.21–0.40 fair, 0.41–0.60 moderate, 0.61–0.80 substantial, 0.81–1.00 excellent.

Results

A total of 112 patients were included. There were 71 men and 41 women. Mean age was 47.1 years (range 21–86)

and mean body mass index was 25.2 ± 3.6 kg/m^2. The fracture was on the left knee in 64 patients (57%) and on the right knee in 48 patients (43%).

While most of the TPF (56%) were sustained during high-energy trauma (traffic accidents, high height falls), in 25% of cases the injury mechanism was a low energy accident and in 19% the mechanism was related to sports injuries.

The inter-observer correlation was fair in the Modified Duparc and Khan classifications. Inter-observer reliability was significantly better in the AO, Schaztker and Luo classifications, with a substantial correlation (Table 1, column A). The intra-observer correlation was moderate in the Modified Duparc and Khan classification and excellent in the AO, Schaztker and Luo classifications (Table 1, column B).

Discussion

One of the most remarkable findings of this investigation was that when combining standard radiographs with 2D and 3D CT images, the AO, Shatzker and Luo classifications showed an excellent intra-observer agreement with a good inter-observer correlation. This was in contrast with our hypothesis. The small discrepancies between intra- and inter-observer agreements might indicate that AO, Schatzker and particularly Luo classifications can be reproducible with adequate training. Zhu et al. [2] also observed a moderate–good inter-observer agreement for these three classification systems. However, they evaluated a considerably lower number of cases (n = 50), and compared only three classifications. Furthermore, the trainers underwent previous training, and although this could theoretically have improved their inter-observer agreement values, it did not positively influence their findings when compared with the present study (Table 2). In another study, Mellema et al. [20] observed only a fair inter-observer agreement in the Shatzker and Luo classifications. They noted that even the addition of 3D CT reconstruction did

Table 2 Published studies

Study group	Fractures	Observers	AO Clasif. Inter / Intra-observer (k)	Schatzker Clasif. Inter / Intra-observer (k)	Luo Clasif. Inter / Intra-observer (k)	Revised Duparc Clasif. Inter / Intra-observer (k)
Hu et al.[a] [14]	21	4	0.86 / 0.89	0.83 / 0.89	-	-
Doornberg et al.[a] [12]	45	6	0.536 / 0.765	0.596 / 0.746	-	-
Gicquel et al. [16]	50	6	0.479 / 0.694	0.476 / 0.660	-	0.474 / 0.784
Zhu et al. [2]	278	4	0.62 / -	0.57 / -	0.77 / -	-
Mellema et al. [20]	15	81	-	0.32 / -	0.28 / -	-
Present study	112	4	0.62 / 0.85	0.65 / 0.87	0.73 / 0.86	0.37 / 0.56

[a] Results using 3D reconstructive evaluation

not improve the overall inter-observer reliability in these cases. Although 81 observers were involved in their study, they only evaluated both these classifications in 15 complex TPF cases (Table 2), randomized to either 2D or 2D and 3D CT evaluations using web-based platforms. Gicquel et al. [16] observed a moderate inter-observer agreement in the 50 cases they evaluated using the Schatzker, AO and Duparc classifications. The more favourable inter-observer agreement observed in our study might be attributed to the fact that we used a combination of standard radiography, 2D CT and 3D CT reconstructions. Furthermore, we provided observers with pre-selected fixed images instead of the whole CT study. The number of evaluators was similar in most studies with the exception of one study [20]. Surprisingly few studies have evaluated intra-observer agreement. In one of these few, Gicquel et al. [16] observed a substantial intra-observer correlation in the classifications of Schatzker, AO and Duparc. They also compared their findings with those from other studies that had calculated this intra-observer correlation, and found that the correlation was better when observers were given 3D CT images [12, 15]. The intra-observer results of the current investigation were also clearly higher (Table 2).

The reproducibility of fracture classifications can be affected by several factors, such as the experience of the observers, the simplicity of the system, additional tools or information provided to the observers, binary decision-making, and rank-order analysis [21]. While intra-observer variability is high, moderate inter-observer agreement could be improved with training and by providing observers with more tools and details to increase accuracy and reproducibility. The use of a combination of standard radiographies, 2D and 3D TC images might be a good alternative to increase the reliability and reproducibility of these frequently used simple classifications. The most frequently used TPF classification was described by Schatzker et al. [3] with standard radiographies only, and showed a low intra and inter-observer agreement [22]. Standard radiographies are often inaccurate and underestimate the extent of displacement and depression of these fractures [23]. Thus, several authors have evaluated their accuracy with CT scan or magnetic resonance imaging and have reported improved results [11, 24–26]. Brunner et al. [27], for example, found that CT scanning could improve the inter-observer and intra-observer reliability in both the AO and Schatzker classifications. In addition, computed tomography is currently considered an essential tool to diagnose and plan tibial plateau surgery [2, 12, 13]. New surgical technique strategies have recently shown that even in simple cases, 3.5-mm locking plates are biomechanically superior to the use of cannulated screws [28–30].

Tibial plateau fractures show great variability in their patterns. In recent years, much focus has been given to the so-called posterior column, which has been shown to have a major influence on surgical planning, fracture reduction accuracy and functional outcomes [31–33]. Luo's classification addresses this column concept, and although it might

excessively simplify the different fracture patterns, it could be used as a complimentary tool to the most extended AO or Schatzker classifications.

The main weakness of the present study was that preselected CT images were provided to the evaluators instead of the whole CT study. This was done to facilitate their assessments but it may have increased intra- and inter-observer agreement as the sample might have underestimated the heterogeneity of the fracture patterns.

Conclusions

Although previous training could be needed, AO, Schatzker and Luo classifications showed a good reproducibility when the tibial plateau fractures were assessed with a combination of standard radiographies and biplanar and 3D CT images. The Modified Duparc and Khan classifications showed lower results, and therefore their use is not recommended.

Compliance with ethical standards

Conflict of interest On behalf of all authors, the corresponding author states that there are no competing interests.

References

1. Agnew SG (1999) Tibial plateau fractures. Oper Tech Orthoped 9: 197–205
2. Zhu Y, Yang G, Luo CF, Smith WR, Hu CF, Gao H, Zhong B, Zeng BF (2012) Computed tomography-based three column classification in tibial plateau fractures: Introduction of its utility and assessment of its reproducibility. J Trauma and Acute Care Surg 73:731–737
3. Schatzker J, McBroom R, Bruce D (1979) The tibial plateau fracture. The Toronto experience 1968—1975. Clin Orthop Relat Res 138:94–104
4. Honkonen SE, Järvinen MJ (1992) Classification of fractures of the tibial condyles. Bone Joint Surg Br 74:840–847
5. Muller ME, Allgower M, Schneider R, Willenegger H (1992) Manual der Osteosynthese. Springer, New York-Berlin-Heidelberg
6. Hohl M (1967) Tibial condylar fractures. J Bone Joint Surg Am 49: 1455–1467
7. Moore TM (1981) Fracture-dislocation of the knee. Clin Orthop Relat Res 156:128–140
8. Wahlquist M, Iaguilli N, Ebraheim N, Levine J (2007) Medial tibial plateau fractures: a new classification system. J Trauma 63:1418–1421
9. Khan RMS, Khan SH, Ahmad AJ, Muhammad U (2000) Tibial plateau fractures: a new classification scheme. Clin Orthop Relat Res 375:231–242
10. Marsh JL, Slongo TF, Agel J, Broderick JS, Creevey W, DeCoster TA et al (2007) Fracture and dislocation classification compendium

- 2007: Orthopaedic Trauma Association classification, database and outcomes committee. J Orthop Trauma 21:S1–33

11. Macarini L, Murrone M, Marini S, Calbi R, Solarino M, Moretti B (2004) Tibial plateau fractures: evaluation with multidetector-CT. Radiol Med 108:503–514

12. Doornberg JN, Rademakers MV, van den Bekerom MP, Kerkhoffs GM, Ahn J, Steller EP et al (2011) Two-dimensional and three-dimensional computed tomography for the classification and characterisation of tibial plateau fractures. Injury 42:1416–1425

13. Charalambous CP, Tryfonidis M, Alvi F, Moran M, Fang C, Samarji R et al (2007) Inter and intra-observer variation of the Schatzker and AO/OTA classifications of tibial plateau fractures and a proposal of a new classification system. Ann R Coll Surg Engl 89:400–404

14. Hu YL, Ye FG, Ji AY, Qiao GX, Liu HF (2009) Three-dimensional computed tomography imaging increases the reliability of classification systems for tibial plateau fractures. Injury 40:1282–1285

15. Muller ME, Nazarian S, Koch P (1987) Classification AO des fractures. Springer, Berlin

16. Gicquel T, Najihi N, Vendeuvre T, Teyssedou S, Gayet LE, Huten D (2013) Tibial plateau fractures: Reproducibility of three classifications (Schatzker, AO, Duparc) and a revised Duparc classification. Orthop Traumatol Surg Res 99:805–881

17. Cohen J (1960) A coefficient of agreement for nominal scales. Educ Psycho Meas 20:37–46

18. Petrie A (2006) Statistics in orthopaedic papers. J Bone Joint Surg (Br) 88:1121–1136

19. Landis JR, Koch GG (1977) The measurement of observer agreement for categorical data. Biometrics 33:159–174

20. Mellema JJ, Doornberg JN, Molenaars RJ, Ring D, Kloen P, Babis GC et al (2016) Interobserver reliability of the Schatzker and Luo classification systems for tibial plateau fractures. Injury 47:944–949

21. Audigé L, Bhandari M, Kellam J (2004) How reliable are reliability studies of fracture classifications? A systematic review of their methodologies. Acta Orthop Scand 75:184–194

22. Te Stroet MAJ, Holla M, Biert J, Van Kampen A (2011) The value of a CT scan compared to plain radiographs for the classification and treatment plan in tibial plateau fractures. Emerg Radiol 18:279–283

23. Dias JJ, Stirling AJ, Finlay DB, Gregg PJ (1987) Computerised axial tomography for tibial plateau fractures. J Bone Joint Surg (Br) 69:84–88

24. Markhardt BK, Gross JM, Monu J (2009) Schatzker classification of tibial plateau fractures: Use of CT and MR imaging improves assessment. Radiographics 29:585–597

25. Wicky S, Blaser PF, Blanc CH, Leyvraz PF, Schnyder P, Meuli RA (2000) Comparison between standard radiography and spiral CT with 3D reconstruction in the evaluation, classification and management of tibial plateau fractures. Eur Radiol 10:1227–1232

26. Yang G, Zhu Y, Luo C, Putnis S (2012) Morphological characteristics of Schatzker type IV tibial plateau fractures: a computer tomography based study. Int Orthop 36:2355–2360

27. Brunner A, Horisberger M, Ulmar B, Hoffmann A, Babst R (2010) Classification systems for tibial plateau fractures; does computed tomography scanning improve their reliability? Injury 41:173–178

28. Carrera I, Gelber PE, Chary G, González-Ballester MA, Monllau JC, Noailly J (2016) Fixation of a split fracture of the lateral tibial plateau with a locking screw plate instead of cannulated screws would allow early weight bearing: a computational exploration. Int Orthop 40:2163–2169

29. Chang SM, Hu SJ, Zhang YQ, Yao MW, Ma Z, Wang X, Dargel J, Eysel P (2014) A surgical protocol for bicondylar four-quadrant tibial plateau fractures. Int Orthop 38:2559–2564

30. Ehlinger M, Adamczewski B, Rahmé M, Adam P, Bonnomet F (2015) Comparison of the pre-shaped anatomical locking plate of 3.5 mm versus 4.5 mm for the treatment of tibial plateau fractures. Int Orthop 39:2465–2471

31. Weil YA, Gardner MJ, Boraiah S, Helfet DL, Lorich DG (2008) Posteromedial supine approach for reduction and fixation of medial and bicondylar tibial plateau fractures. J Orthop Trauma 22:357–362

32. Higgins TF, Kemper D, Klatt J (2009) Incidence and morphology of the posteromedial fragment in bicondylar tibial plateau fractures. J Orthop Trauma 23:45–51

33. Zeng ZM, Luo CF, Putnis S, Zeng BF (2011) Biomechanical analysis of posteromedial tibial plateau split fracture fixation. Knee 18:51–54