

ANALYSIS OF THE REPRODUCIBILITY OF TIBIAL PLATEAU FRACTURES' CLASSIFICATION

Rodrigo Pires e Albuquerque¹, Vincenzo Giordano², Alexandre Pallottino³, Tannous Sassine⁴, Rodrigo Canedo⁴, Juliana Pina⁴, Ney Pecegueiro do Amaral⁵

ABSTRACT

Objective: To assess the inter-observer agreement of three tibial plateau fracture classifications: Schatzker, AO Group, and Hohl. **Methods:** Nine physicians of different levels of expertise (six Orthopaedic Residents – OR – and three Orthopaedic Attendings – OA) classified 50 tibial plateau fractures. **Results:** There was a low to moderate agreement between OR and OA on the three classification systems. A straight correlation was found between the year of Residence and an increase on the level of agreement on the AO group and Hohl classifications.

This was partially seen on the Schatzker classification. Kappa ranged from 0.344 to 0.577, 0.36 to 0.499, and 0.278 to 0.465 for the Hohl, AO group, and Schatzker systems, respectively. **Conclusion:** With a level of 0.1, there was a significant agreement on the three classifications among the physicians. When comparing ORs to OAs, the three studied classifications show low to moderate consistency, with the Hohl classification showing the highest level of agreement between OR.

Keywords - Tibial fractures/classification; Validation studies; Evaluation

INTRODUCTION

Tibial plateau fractures account for 1-2% of all fractures and about 8% of fractures in the elderly⁽¹⁾. Because they are fractures that involve articular cartilage, they require anatomical reduction and early mobilization, which are preferably obtained by surgical means⁽²⁾. In this sense, the use of a comprehensive classification system makes the definition of tactics and surgical access more understandable.

Conceptually, the proposal to adopt a classification system happens due to its reproducibility, simplicity, and ease of memorization, aiding in the choice of treatment and prognostic prediction of the existing injury. In tibial plateau fractures, the most commonly used classifications are the Schatzker⁽³⁾, the AO Group's (Arbeitsgemeinschaft für Osteosynthesefragen)⁽⁴⁾ and that of Hohl⁽⁵⁾.

The objective of this study was to evaluate inter-

observer agreement in the Schatzker, AO Group, and Hohl classifications in a group of doctors divided by different degrees of professional experience.

METHOD

A prospective study of the reproducibility of the most widely used classifications for fractures of the tibial plateau was conducted between the months of October 2005 and June 2006 at the Professor Nova Monteiro Orthopedics and Traumatology Clinic, Miguel Couto Municipal Hospital (HMMC, Hospital Municipal Miguel Couto). We evaluated the knee radiographs of 50 subjects diagnosed with tibial plateau fracture treated in the HMMC. All were skeletally mature, with a mean age of 47.9 years, ranging from 21 to 89 years; 15 were women and 35 were men. Twenty-one fractures occurred on the right side and 29 on the left. Anteroposterior and lateral radiographs of the knee

1 – Orthopedist; Coordinator, Knee Surgery Department, Orthopedics and Traumatology Clinic, Hospital Municipal Miguel Couto (SOT-HMMC), Rio de Janeiro, Brazil.

2 – Orthopedist; Head, Medical Residency Program, SOT-HMMC, RJ, Brazil.

3 – Orthopedist; Assistant, SOT-HMMC, RJ, Brazil.

4 – Orthopedist; SOT-HMMC, RJ, Brazil.

5 – Head, SOT-HMMC, RJ, Brazil.

Study conducted at the Professor Nova Monteiro Orthopedics and Traumatology Clinic (Serviço de Ortopedia e Traumatologia Professor Nova Monteiro) – Hospital Municipal Miguel Couto (SOT-HMMC), Rio de Janeiro, RJ, Brazil.

Correspondence: Av. Henrique Dodsworth, 83, apto. 105, CEP: 22061-030, Rio de Janeiro, RJ, Brasil. Tel/Fax: (21) 2287-2063. E-mail: rodalbuquerque@ibest.com.br

We declare no conflict of interest in this article

on the injury side were obtained from all patients, following standard protocol. Patients were positioned supine with their knee in maximum extension and with the leg in a neutral position. The X-ray tube was placed 1 m away from the injured site with the radius directed orthogonally to the inferior pole of the patella at 5° cephalic angulation. In the lateral radiograph, the patient maintained the knee at 20° of flexion with the tube-film distance standardized at 1 m and at 5° cephalic angulation. The Super 100® X-ray machine (Philips, Brazil) was used at 50kV and 31mA. The results were evaluated by the researchers as to the quality of the image and the exam was repeated if it was judged to be of poor technical quality.

The tests were shown to a group of nine physicians with varying degrees of experience in orthopedics and traumatology. The group was composed of three medical members of the Brazilian Society of Orthopedic Trauma (Sociedade Brasileira de Trauma Ortopédico), considered medical specialists (MS), and six medical residents (MR), two of whom were in their third year (R3), two in their second year (R2) and two in their first year (R1). All participants were informed of the purpose of this investigation. The project was submitted for approval by the Commission for Research Ethics of the Municipal Health Secretariat of Rio de Janeiro, according to Resolution 196/96 of the National Health Council (Guidelines and Norms Regulating Research Involving Human Subjects)⁽⁶⁾. The patients were informed of the study and the patient or a close relative signed a consent form.

In order to minimize bias due to the difficulty of interpretation or any possible instance of forgetting, the classifications were described on the answer sheet with associated schematic drawings of the respective classifications given to each observer upon evaluation of the radiographs (Figures 1, 2, and 3). There was no time limit for classifying the radiographs.

The radiographic analysis data was collected blindly by one of the authors, who collected the forms and scanned the responses of the nine participants. Statistical analysis was performed using the kappa statistic, which evaluates interobserver agreement for qualitative data. The hypothesis tested is whether the agreement (p) is equal to zero, that is, that there is no interobserver agreement. It is known that the kappa statistic ranges from a negative value to 1.0 (kappa = 1 expresses perfect agreement, kappa close

to zero expresses disagreement, i.e., that the observed agreement is no better than chance). An α of 1% was adopted as the level of significance (Figures 1, 2, 3).

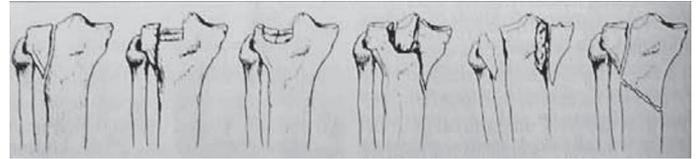


Figure 1 – Schatzker classification.

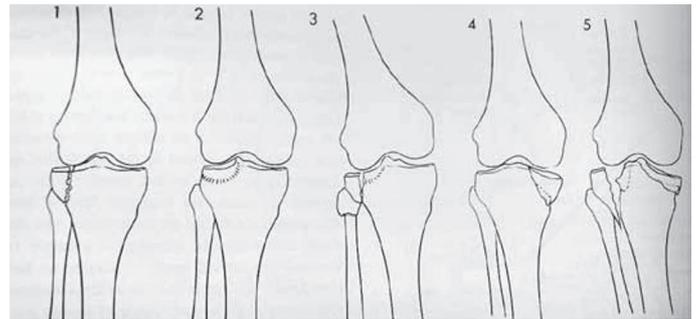


Figure 2 – Hohl classification.

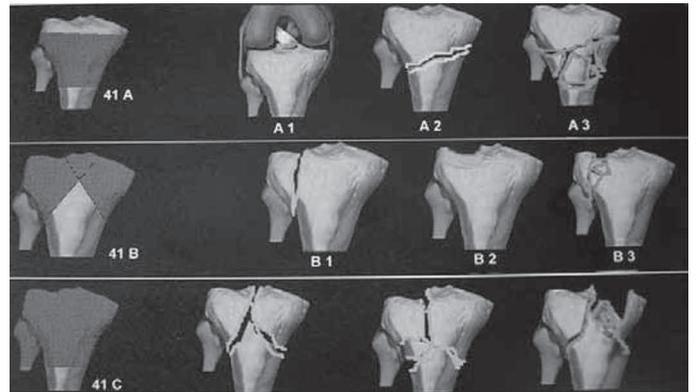


Figure 3 – AO/ASIF classification.

RESULTS

Tables 1, 2, 3, and 4 reveal the frequency (n) and percentage (%) of the Schatzker, AO, and the Hohl classifications of the nine observers for the 50 patients studied.

Tables 5, 6, and 7 provide the agreement observed (in %), kappa, kappa standard error (SE), and the descriptive level (p-value) for each pair of observers considered for the Schatzker, AO, and Hohl classifications.

The agreement observed corresponds to the percentage (%) of identical answers in the 50 patients total. For example, in the 50 patients total, according to the Schatzker classification, the R3_1 agreed with the MS on 58% of cases.

Table 1 – Descriptive analysis for the Schatzker classification.

Classification	MS		R3_1		R3_2		R2_1		R2_2		R1_1		R1_2	
	n	%	n	%	n	%	n	%	n	%	n	%	n	%
Schatzker 1	7	14	10	20	6	12	11	22	10	20	14	28	11	22
Schatzker 2	21	42	16	32	10	20	15	30	21	42	17	34	17	34
Schatzker 3	7	14	7	14	18	36	6	12	3	6	6	12	4	8
Schatzker 4	2	4	5	10	3	6	7	14	4	8	3	6	6	12
Schatzker 5	8	16	4	8	4	8	5	10	2	4	3	6	5	10
Schatzker 6	5	10	8	16	9	18	6	12	10	20	7	14	7	14

MS: medical specialists, R3: third-year resident, R2: second-year resident, R1: first-year resident.
Source: SOT-HMMC, 2006.

Table 2 – Descriptive analysis for the AO Group classification.

Classification	MS		R3_1		R3_2		R2_1		R2_2		R1_1		R1_2	
	n	%	n	%	n	%	n	%	n	%	n	%	n	%
AO A1	0	0	0	0	1	2	0	0	1	2	1	2	1	2
AO A2	0	0	1	2	1	2	0	0	1	2	0	0	2	4
AO B1	9	18	10	20	6	12	14	28	14	28	14	28	12	24
AO B2	8	16	10	20	20	40	9	18	8	16	7	14	10	20
AO B3	21	42	17	34	12	24	16	32	16	32	18	36	14	28
AO C1	7	14	8	16	5	10	4	8	4	8	6	12	5	10
AO C2	2	4	1	2	2	4	4	8	2	4	2	4	0	0
AO C3	3	6	3	6	3	6	3	6	4	8	2	4	6	12

MS: medical specialists, R3: third-year resident, R2: second-year resident, R1: first-year resident.
Source: SOT-HMMC, 2006.

Table 3 – Descriptive analysis for the pooled AO Group classification.

Classification	MS		R3_1		R3_2		R2_1		R2_2		R1_1		R1_2	
	n	%	n	%	n	%	n	%	n	%	n	%	n	%
AO A1/A2	0	0	1	2	2	4	0	0	2	4	1	2	3	6
AO B1	9	18	10	20	6	12	14	28	14	28	14	28	12	24
AO B2	8	16	10	20	20	40	9	18	8	16	7	14	10	20
AO B3	21	42	17	34	12	24	16	32	16	32	18	36	14	28
AO C1	7	14	8	16	5	10	4	8	4	8	6	12	5	10
AO C2/C3	5	10	4	8	5	10	7	14	6	12	4	8	6	12

MS: medical specialists, R3: third-year resident, R2: second-year resident, R1: first-year resident.
Source: SOT-HMMC, 2006.

Table 4 – Descriptive analysis for the Hohl classification.

Classification	MS		R3_1		R3_2		R2_1		R2_2		R1_1		R1_2	
	n	%	n	%	n	%	n	%	n	%	n	%	n	%
Hohl 1	7	14	10	20	6	12	11	22	9	18	15	30	11	22
Hohl 2	9	18	6	12	18	36	6	12	5	10	5	10	1	2
Hohl 3	20	40	16	32	12	24	16	32	22	44	20	40	21	42
Hohl 4	2	4	5	10	1	2	7	14	4	8	1	2	4	8
Hohl 5	12	24	13	26	13	26	10	20	10	20	9	18	13	26

MS: medical specialists, R3: third-year resident, R2: second-year resident, R1: first-year resident.
Source: SOT-HMMC, 2006.

Table 5 – Interobserver analysis for the Schatzker classification.

Obs 1 x Obs 2	Observed agreement	Kappa coefficient	Kappa SE	p-value
MS x R3_1	58.0	0.465	0.067	< 0.0001
MS x R3_2	52.0	0.411	0.062	< 0.0001
MS x R2_1	48.0	0.344	0.066	< 0.0001
MS x R2_2	50.0	0.340	0.068	< 0.0001
MS x R1_1	44.0	0.278	0.068	< 0.0001
MS x R1_2	54.0	0.411	0.067	< 0.0001
R3_1 x R3_2	52.0	0.415	0.063	< 0.0001
R3_1 x R2_1	54.0	0.426	0.067	< 0.0001
R3_1 x R2_2	64.0	0.535	0.070	< 0.0001
R3_1 x R1_1	54.0	0.414	0.070	< 0.0001
R3_1 x R1_2	56.0	0.446	0.068	< 0.0001
R3_2 x R2_1	54.0	0.447	0.060	< 0.0001
R3_2 x R2_2	52.0	0.419	0.059	< 0.0001
R3_2 x R1_1	54.0	0.440	0.062	< 0.0001
R3_2 x R1_2	40.0	0.283	0.059	< 0.0001
R2_1 x R2_2	54.0	0.413	0.068	< 0.0001
R2_1 x R1_1	50.0	0.368	0.068	< 0.0001
R2_1 x R1_2	44.0	0.297	0.068	< 0.0001
R2_2 x R1_1	60.0	0.473	0.073	< 0.0001
R2_2 x R1_2	62.0	0.504	0.071	< 0.0001
R1_1 x R1_2	66.0	0.564	0.070	< 0.0001

Obs: observer; MS: medical specialists, R3: third-year resident, R2: second-year resident; R1: first-year resident, SE: standard error.

Source: SOT-HMMC, 2006.

Table 6 – Interobserver analysis for the AO Group classification.

Obs 1 x Obs 2	Observed agreement	Kappa coefficient	Kappa SE	p-value
MS x R3_1	62.0	0.499	0.074	< 0.0001
MS x R3_2	58.0	0.468	0.067	< 0.0001
MS x R2_1	62.0	0.501	0.074	< 0.0001
MS x R2_2	52.0	0.374	0.072	< 0.0001
MS x R1_1	54.0	0.388	0.075	< 0.0001
MS x R1_2	50.0	0.360	0.070	< 0.0001
R3_1 x R3_2	64.0	0.544	0.068	< 0.0001
R3_1 x R2_1	60.0	0.484	0.072	< 0.0001
R3_1 x R2_2	56.0	0.436	0.071	< 0.0001
R3_1 x R1_1	64.0	0.531	0.073	< 0.0001
R3_1 x R1_2	54.0	0.418	0.070	< 0.0001
R3_2 x R2_1	56.0	0.447	0.067	< 0.0001
R3_2 x R2_2	54.0	0.428	0.065	< 0.0001
R3_2 x R1_1	54.0	0.427	0.065	< 0.0001
R3_2 x R1_2	40.0	0.250	0.067	< 0.0001
R2_1 x R2_2	56.0	0.426	0.074	< 0.0001
R2_1 x R1_1	52.0	0.369	0.075	< 0.0001
R2_1 x R1_2	48.0	0.335	0.072	< 0.0001
R2_2 x R1_1	58.0	0.450	0.074	< 0.0001
R2_2 x R1_2	56.0	0.440	0.071	< 0.0001
R1_1 x R1_2	66.0	0.565	0.071	< 0.0001

Obs: observer; MS: medical specialists, R3: third-year resident, R2: second-year resident; R1: first-year resident, SE: standard error.

Source: SOT-HMMC, 2006.

Table 7 – Interobserver analysis for the Hohl classification.

Obs 1 x Obs 2	Observed agreement	Kappa coefficient	Kappa SE	p-value
MS x R3_1	68.0	0.577	0.076	< 0.0001
MS x R3_2	64.0	0.526	0.076	< 0.0001
MS x R2_1	60.0	0.478	0.073	< 0.0001
MS x R2_2	60.0	0.452	0.078	< 0.0001
MS x R1_1	52.0	0.348	0.078	< 0.0001
MS x R1_2	52.0	0.344	0.077	< 0.0001
R3_1 x R3_2	60.0	0.491	0.070	< 0.0001
R3_1 x R2_1	62.0	0.509	0.074	< 0.0001
R3_1 x R2_2	66.0	0.547	0.076	< 0.0001
R3_1 x R1_1	58.0	0.441	0.076	< 0.0001
R3_1 x R1_2	60.0	0.462	0.078	< 0.0001
R3_2 x R2_1	54.0	0.424	0.067	< 0.0001
R3_2 x R2_2	54.0	0.413	0.069	< 0.0001
R3_2 x R1_1	54.0	0.414	0.069	< 0.0001
R3_2 x R1_2	42.0	0.272	0.064	< 0.0001
R2_1 x R2_2	54.0	0.392	0.075	< 0.0001
R2_1 x R1_1	50.0	0.338	0.075	< 0.0001
R2_1 x R1_2	52.0	0.361	0.076	< 0.0001
R2_2 x R1_1	60.0	0.446	0.080	< 0.0001
R2_2 x R1_2	60.0	0.441	0.081	< 0.0001
R1_1 x R1_2	66.0	0.525	0.082	< 0.0001

Obs: observer; MS: medical specialists, R3: third-year resident, R2: second-year resident, R1: first-year resident, SE: standard error.
Source: SOT-HMMC, 2006.

The AO Group classification was analyzed by pooling results (Table 3) due to the very small number of cases observed in the A1 and A2 categories, and C1, C2, and C3.

When the MR are compared with the MS, the three classifications showed weak to moderate agreement; none showed good quality.

The AO Group and the Hohl classifications showed an increase in the degree of agreement as the years of residency increased. The same outcome was partially observed for the Schatzker classification.

Based on the above comment, the Hohl classification achieved better results (kappa ranging from 0.344 to 0.577) than the AO Group classification (kappa ranging from 0.36 to 0.499) and the Schatzker classification (kappa ranging from 0.278 to 0.465).

DISCUSSION

In orthopedics and traumatology, classifications are tools that help to standardize an international language for addressing injuries. What usually happens is that once a classification is established for a particular injury, which is not always based on a relevant and

representative sample, some cases do not fit the described or classified types. Weber, in his classic study of malleolar fractures, reserved an “unclassifiable” subgroup of injuries for those that did not fit into classes or groups because of their peculiarity⁽⁷⁾. Some classifications were replaced by other more complete classifications over time.

Even today, there are few clinics in Brazil that have routine access to computed tomography or magnetic resonance imaging, hence the need for a study assessing the radiographic classifications of tibial plateau fractures. Dirschl and Patrick⁽⁸⁾ found that the classifications for fractures of the tibial plateau have low interobserver reproducibility, prompting our interest in this study.

Walton et al.⁽⁹⁾ assessed 53 radiographs of the knee in patients with tibial plateau fractures and compared Schatzker classification with the AO Group classification. They concluded that the AO classification was more reproducible than the Schatzker classification in the interobserver analysis, corroborating our study. They observed that both classifications were originally based on radiographic studies, demonstrating the importance of our study. Charalambous et al.⁽¹⁰⁾ observed that the Schatzker and AO classifications were less reproducible in their intra- and interobserver analysis, confirming the results of our study.

Mandarino et al.⁽¹¹⁾ evaluated 20 radiographs of the knee in patients with tibial plateau fractures in a retrospective study and analyzed the reproducibility of the Schatzker classification. They concluded that the Schatzker classification has moderate interobserver reproducibility in the analysis, even after grouping them into more homogeneous groups, demonstrating what we noted in our study.

Martin et al.⁽¹²⁾ assessed the interobserver reproducibility of the AO classification of the radiographs of 56 fractures of the tibial plateau. They concluded that the AO classification of tibial plateau fractures is reproducible and that agreement on conventional radiographic evaluation is dependent on the observer’s experience. When correlated with our study, we observe that as we increase the degree of experience of the physicians, we increase agreement, revalidating our research.

Khan et al.⁽¹³⁾ developed a new classification for the evaluation of tibial plateau fractures. Eighty patients with tibial plateau fractures were evaluated

retrospectively. They observed that the existing classifications had flaws and sometimes different fractures were graded into the same subgroup. When we correlated their results with our work, we observed difficulty in grading fractures with the posterior fragment of the tibial plateau into existing classifications. Rafii et al.⁽¹⁴⁾ demonstrated the superiority of computed tomography (CT) over conventional radiography in fractures of the tibial plateau. CT, the authors say, is a reliable method for the evaluation and classification of tibial plateau fractures. Nevertheless, the reality is that, in most hospitals, CT is not available, in addition to being costly and requiring a high degree of exposure to radiation.

Maripuri et al.⁽¹⁵⁾ conducted a study similar to ours in which they compared the Schatzker, AO, and Hohl classifications. They noted that none of these classifications is ideal. When we analyzed our study, we found that there is a weak to moderate agreement.

CONCLUSIONS

There is significant agreement at the 1% level among the three classifications for all pairs of observers. Comparing the MR with the MS, the three classifications studied showed weak and moderate agreement. The Hohl classification showed the highest agreement between the MR.

REFERENCES

- Jacofsky DJ, Haidukerwych GJ. Tibia plateau fractures. In: Scott WN. Insall & Scott surgery of the knee. Philadelphia: Churchill Livingstone; 2006. p.1133-46.
- Stover MD, Kellam JF. Fraturas articulares: princípios. In: Ruëdi TP, Murphy WM. Princípios AO do tratamento de fraturas. Porto Alegre: Artmed; 2002. p.105-19.
- Schatzker J, Mc Broom R, Bruce D. The tibial plateau fracture: the Toronto experience 1968-1975. Clin Orthop Relat Res. 1979;(138):94-104.
- Murphy WM, Leu D. Fraturas articulares: princípios. In: Ruëdi TP, Murphy WM. Princípios AO do tratamento de fraturas. Porto Alegre: Artmed; 2002. p.45-58
- Hohl M, Luck J. Fractures of the tibial condyle: a clinical and experimental study. J Bone Joint Surg Am. 1956;38(5):1001-18.
- Comissão Nacional de Ética em Pesquisa (CONEP). <http://conselho.saúde.gov.br/comissão/eticapesq.htm>. Acessado em dezembro de 2006.
- Franco JS, Musafir M. Classificação das fraturas mais freqüente. Rio de Janeiro: Schering- Plough; 1991.
- Dirschl DR, Patrick DA. Injury severity assessment in tibial plateau fractures. Clin Orthop Relat Res. 2004;(423):85-92.
- Walton NP, Harish S, Roberts C, Blundell C. AO or Schatzker? How reliable is classification of tibial plateau fractures? Arch Orthop Trauma Surg. 2003; 123(8):396-8.
- Charalambous CP, Tryfonidis M, Alvi F, Moran M, Fang C, Samaraji R, Hirst P. 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. 2007;89(4):400-4.
- Mandarino M, Pessoa A, Guimarães JAM. Avaliação da reprodutibilidade da classificação de Schatzker para as fraturas do planalto tibial. R INTO. 2004; 2(2):11-8.
- Martin J, Marsh JL, Nepola JV, Dirschl DR, Hurwitz S, DeCoster TA. Radiographic fracture assessments: Which ones can we reliably make? J Orthop Trauma. 2000;14(6):379-85.
- Khan RMS, Khan SH, Ahmad AJ, Umar M. Tibial Plateau Fractures: A New Classification Scheme. Clin Orthop Relat Res. 2000;(375):231-42.
- Rafii M, Lamont JG, Firooznia H. Tibial plateau fractures: CT evaluation and classification. Crit Rev Diagn Imaging. 1987;27(2):91-112.
- Maripuri SN, Rao P, Manoj-Thomas A, Mohanty K. The classifications systems for tibial plateau fractures: How reliable are they? Injury. [serial on the internet]. 2008. Available from ://www.elsevier.com/locate/injury