## **OPEN** Research Article

Intraobserver and Interobserver Reproducibility of Classifications of Tibial Plateau Fractures and the Surgical Approaches Chosen Comparing 2D CT and 3D Printing: Reliability Study

Jose Osma-Rueda, MD, MSc D Fabian Mantilla-Angarita, MD Clara Lopez-Gualdron, PhD

From the Faculty of Health Sciences, Universidad Industrial de Santander, Sociedad Colombiana de Cirugía Ortopédica y Traumatología (SCCOT), Bucaramanga, Colombia (Dr. Osma and Dr. Mantilla), and the Industrial Design College, Universidad Industrial de Santander, Bucaramanga, Colombia (Dr. López).

Correspondence to Dr. Osma: osmarueda@gmail.com

None of the following authors or any immediate family member has received anything of value from or has stock or stock options held in a commercial company or institution related directly or indirectly to the subject of this article: Dr. Osma, Dr. Mantilla, and Dr. López.

JAAOS Glob Res Rev 2023;7: e22.00202

DOI: 10.5435/JAAOSGlobal-D-22-00202

Copyright 2023 The Authors. Published by Wolters Kluwer Health, Inc. on behalf of the American Academy of Orthopaedic Surgeons. This is an open access article distributed under the Creative Commons Attribution License 4.0 (CCBY), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

### ABSTRACT

**Introduction:** Reliability is the study of internal consistency, reproducibility (intraobserver and interobserver), and agreement. Reproducibility studies that classify tibial plateau fractures have used plain radiography and two-dimensional (2D) CT scans and threedimensional (3D) printing. The objective of this study was to evaluate the reproducibility of the Luo Classification of tibial plateau factures and the surgical approaches chosen for these fractures based on 2D CT scans and 3D printing.

**Methods:** A reliability study was performed at the Universidad Industrial de Santander, Colombia, that evaluated the reproducibility of the Luo Classification of tibial plateau fractures and the choice of surgical approaches based on 20 CT scans and 3D printing, with five evaluators. **Results:** For the trauma surgeon, reproducibility was better when evaluating the classification using 3D printing, with a kappa of 0.81 (95% confidence interval [CI], 0.75-0.93; P < 0.01) than when using CT scans, with a kappa of 0.76 (95% CI, 0.62-0.82; P < 0.01). When comparing the surgical decisions made by the fourth-year resident with those of the trauma surgeon, a fair reproducibility was obtained using CT, with a kappa of 0.34 (95% CI, 0.21-0.46; P < 0.01), which improved to substantial when using 3D printing, with a kappa of 0.63 (95% CI, 0.53-0.73; P < 0.01).

**Discussion:** This study found that 3D printing provided more information than CT and decreased measurement errors, thereby improving reproducibility, as shown by the higher kappa values that were obtained.

**Conclusion:** The use of 3D printing and its usefulness are helpful to decision making when providing emergency trauma services to patients with intraarticular fractures such as those of the tibial plateau.

ibial plateau fractures of the knee represent approximately 1% of all fractures<sup>1</sup> and 8% of all fractures in adults.<sup>2</sup> Fractures on the surface of the tibial plateau present the risk of biomechanical instability of the knee articulation.<sup>3,4</sup> Therefore, a diagnosis that helps to make decisions about the effective treatment of these fractures is crucial.<sup>5</sup> Nevertheless, when complex intraarticular fracture patterns occur in the area of the tibial plateau, it is not always easy to identify the fracture lines.<sup>6</sup> This can result in performing diagnostics and/or surgical approaches that are not indicated. It is important that orthopedic and trauma surgeons have better information for making decisions, which can come from threedimensional (3D) printing models.

An entirely reliable classification for this type of fracture does not currently exist, which is why we suggest that the use of 3D printing can decrease measurement errors in reliability.

Classifications of fractures should be validated, including tibial plateau fractures. This includes content and construct validation.<sup>7-9</sup>

Reliability is studied based on three factors: internal consistency, reproducibility, and agreement. Intraobserver and interobserver reliability of the classification of tibial plateau fractures (reproducibility) has used plain radiography, two-dimensional (2D) CT scans, and 3D reconstruction, with increasingly better reliability values being obtained since the technological evolution of the imaging has improved the information that the evaluator and the clinician depend on to make decisions. This has resulted in better clinical outcomes for patients and improved reliability.<sup>5,7,9</sup>

With the technological advances that have been made in medicine, some studies have combined imaging techniques such as CT and new technologies such as 3D printing for the reconstruction of anatomical models.<sup>10-12</sup> Just as with the evolution of information from diagnostic images, 3D printing provides better and more real information, which helps to reach more precise diagnoses based on precise anatomical models.<sup>12</sup> As shown by prior evidence, 3D printing can be used to visualize the characteristics of tibial plateau fractures on axial, sagittal, and coronal planes.

The objective of this study was to evaluate intraobserver and interobserver reliability of the Luo Classification of tibial plateau fractures<sup>13</sup> and the choice of surgical approaches for these fractures by comparing 2D CT scans with 3D printing.

### Methods

With approval from the Ethics and Research Committee of the Universidad industrial de Santander and the Hospital Universitario de Santander in Bucaramanga, Colombia, an interobserver and intraobserver reliability study was performed between the years 2017 and 2019 to assess reproducibility when evaluating the Luo Classification<sup>13</sup> of tibial plateau fractures and making decisions about the surgical approaches for each of the cases evaluated (Figure 1). For this purpose, the study used CT scans of different types of tibial plateau fractures from 20 patients not randomized. The inclusion criteria were patients older than 18 years of age who had an unintentional injury, as injuries caused by transit or injuries caused by dairy activities. Patients with pathological fractures, injuries caused by gun weapons, blunt trauma, and open fractures were excluded.

The images of the fractures were obtained with CT image segmentation. The CT scans were generated with an Aquilion 64-slice helical system with 0.5 mm slices of the tibial plateau and were stored with digital imaging and communications in medicine. The 3D printing models of the 20 cases shown in Figure 2 were obtained with free reverse engineering software known as In-Vesalius,<sup>14</sup> which specializes in generating virtual 3D imaging. These were prepared with computer-aided design software, after which 3D models were printed at actual scale in polylactic acid using fused deposition modeling, with a resolution of 0.1 mm. The Interfaz group conducted the process for generating the biomodels at the 3D technology laboratory at the Universidad Industrial de Santander.<sup>15,16</sup>

Five evaluators were included in this study: three orthopaedic surgeons and traumatologists, one specialized in trauma surgery, one knee surgeon, and one general orthopaedic surgeon; and two students of a third-year and a fourth-year resident of orthopedy and traumatology who had previously reviewed the Luo Classification<sup>13</sup> and the surgical approaches, namely anterolateral, anteromedial, midline, posteromedial, posterolateral, posterior inverted L, and minimally invasive. A pilot test was performed with three cases to become familiar with the data collection instrument, the classification, and the surgical approaches used for tibial plateau fractures.

The 5 evaluators assessed the 20 cases blindly, independently, at different times, and randomized. The classifications and choice of surgical approaches were evaluated for each case according to the type of fracture, based on the Luo Classification of the four tibial plateau columns (Figure 1).<sup>13</sup> Each evaluator assessed the 2D CT and 3D printing of each case under the same conditions, every 2 weeks, but the durations of the evaluation time was not standardized.

### Figure 1



Illustration showing the classification by Luo et al. of the plateu fractures.<sup>13</sup> AL = Anterolateral, AM = Anteromedial, PL = Posterolateral, PM = Posteromedial.

# Stages and Times for Evaluating the 2D CT Scans and 3D Printing Models

Next, the activities that were performed will be described and the time during the study when they were conducted, with a difference of 2 weeks between the first and second assessment for CT and 3D, evaluating classification and surgical approach.

### First

The CT images that were collected and printed of the tibial plateaus were reviewed to evaluate the classifications.

### Second

Two weeks after the first evaluation, the 3D printing models were reviewed to evaluate the classifications.

### Third

Four weeks after the first evaluation, another assessment was performed of the CT scan images to choose the surgical approaches of the tibial plateau fractures.

### Fourth

Six weeks after the first evaluation, a second assessment was performed of the classifications and choice of surgical approaches based on each of the 3D printing models of the tibial plateau fractures.

## Information Processing and Analysis Statistics

The information obtained with the data collection instruments was entered in duplicate in an Excel spreadsheet, and the data were processed using STATA version 16 software. A descriptive analysis was performed of the sociodemographic characteristics of the patients included in the study and the frequency distributions. The continuous variables according to the number of patients included (n = 20) and the distribution are presented in medians and interquartile ranges.

Reproducibility was evaluated with the kappa statistic and their 95% confidence interval (CI),<sup>17,18</sup> which were interpreted based on the proposal by Landis and Koch<sup>19</sup>: 0.01 to 0.20 slight, 0.21 to 0.40 fair, 0.41 to 0.60 moderate, 0.61 to 0.80 substantial and 0.81 to 1.00 almost perfect.

### **Results**

A total of 200 evaluations were performed of intraobserver reliability for each evaluation method: 2D CT and 3D printing. Another 200 evaluations were conducted of interobserver reliability for the surgical approaches that were selected for the different types of tibial plateau fractures based on the CT scans and 3D printing.

### Sociodemographic Characteristics

In the main overall characteristics of the 20 patients in the study, the median age was 34 years (18-78) and most were men (18 of 20). The left side was most often affected, with 11 patients, and the most frequent mechanism of injury was trauma caused by traffic accidents (17 patients), with motorcycle crashes being the most common (13 patients).

## Intraobserver Reproducibility of the Classification With CT Scans

In the intraobserver reproducibility of the Luo Classification of the tibial plateau fractures when using CT

### Figure 2



Illustration showing the 3D-printing biomodels used in the study.

scans, based on Landis and Koch,<sup>19</sup> the assessment by the trauma surgeon was substantial, with a kappa of 0.76 (95% CI, 0.62-0.82; P < 0.01), followed by the knee surgeon and the fourth-year resident who obtained a kappa of 0.69 (95% CI, 0.60-0.80; P < 0.01) (Table 2). In the surgical approaches chosen based on reviewing the CT scans, the intraobserver reproducibility for the trauma surgeon was almost perfect, with a kappa of 0.69 (95% CI, 0.60-0.80; P < 0.01) and 85% agreement, followed by the knee surgeon with a substantial kappa of 0.67 (95% CI, 0.57-0.76; P < 0.01) and 75% agreement (Table 1).

### Intraobserver Reproducibility of the Classification With 3D Printing

In respect of the classification of tibial plateau fractures, intraobserver reproducibility with 3D printing was almost perfect for the trauma surgeon, with a kappa of 0.81 (95% CI, 0.75-0.93; P < 0.01) (Table 1).

### Intraobserver Reproducibility of Classifications and Surgical Approaches Using 3D Printing

Furthermore, reproducibility for the group of evaluators was better with 3D printing than with 2D CT scans. Reproducibility of the choice of surgical approach was also better with 3D printing for the entire group of evaluators, and the trauma surgeon continued to be better, having obtained an almost perfect result with a kappa of 0.93 (95% CI, 0.92-1.00; P < 0.01) (Table 1).

### Interobserver Reproducibility of Classifications and Surgical Approaches Using 3D Printing

When using 3D printing, interobserver reliability was slight for the classification, with a kappa of 0.18 (95% CI, 0.12-0.19), and moderate for the surgical approach, with a kappa of 0.47 (95% CI, 0.36-0.56). Both were statistically significant (P < 0.01).

Finally, when analyzing the interobserver reproducibility of the surgical approach by comparing the more experienced trauma surgeon with the other evaluators, reproducibility was better with the 3D printing for both the Luo Classification<sup>13</sup> and the surgical approach (Table 2).

### **Discussion**

This reliability study compared the interobserver and intraevaluator reproducibility of the Luo Classifications of tibial plateau fractures and the corresponding choice of surgical approaches based on 3D printing versus 2D CT scans. When comparing 3D printing and 2D CT scans, the intraobserver reproducibility of the classification and surgical approach was always better with 3D printing for each of the five evaluators, with the surgical approach being even better and statistically significant. In addition, the most experienced evaluator obtained an almost perfect kappa value<sup>19</sup> for the surgical approach when using 3D printing models. When comparing 2D CT scans and 3D printing for the fourth-year resident, intraobserver

**Research** Article

	CT/Classification			3D/Classification			CT/Approach			3D/Approach		
Evaluator	Kappa	95% CI	%	Kappa	95% CI	%	Kappa	95% CI	%	Kappa	95% CI	%
Trauma surgeon	0.76 <sup>a</sup>	0.62-0.82	80	0.81 <sup>a</sup>	0.75-0.93	85	0.81 <sup>a</sup>	0.79-0.93	85	0.93 <sup>a</sup>	0.91-1.00	95
Knee surgeon	0.69 <sup>a</sup>	0.60-0.80	75	0.53 <sup>a</sup>	0.47-0.64	60	0.67 <sup>a</sup>	0.59-0.76	75	0.58 <sup>a</sup>	0.41-0.70	65
Orthopaedic surgeon	0.29 <sup>a</sup>	0.16-0.37	45	0.49 <sup>a</sup>	0.41-0.53.	60	0.59 <sup>a</sup>	0.37-0.63	75	0.76 <sup>a</sup>	0.38-0.92.	85
Third-year resident	0.55 <sup>a</sup>	0.43-0.71	60	0.50 <sup>a</sup>	0.36-0.60.	60	0.50 <sup>a</sup>	0.25-0.82	55	0.51 <sup>a</sup>	0.45-0.55	65
Fourth-year resident	0.65 <sup>a</sup>	0.50-0.73	70	0.27 <sup>a</sup>	0.07-0.38	35	0.54 <sup>a</sup>	0.38-0.62	65	0.80 <sup>a</sup>	0.55-0.92	85

Table 1. Intraobse	erver Reliability With 2D	CT: 3D Printing for the Luo	Classification and Surgical Approach
--------------------	---------------------------	-----------------------------	--------------------------------------

2D = two-dimensional, 3D = three-dimensional printing, CI = confidence interval  ${}^{a}\!P < 0.01.$ 

reliability of the surgical approach improved from a kappa of 0.54 (95% CI, 0.38-0.62) with the CT scans to a kappa of 0.80 (95% CI, 0.55-0.92) with 3D printing (Table 1).

Regarding interobserver reproducibility, the kappa value was found to be affected by the prevalence of the event and by bias because of experience (paradox bias),<sup>9</sup> which in this study related to the trauma surgeon who had the most experience and whose evaluations were more reliable than those of the two residents in training. The lack of experience by the residents resulted in slight interobserver reliability values (Table 2).<sup>18</sup> Under the scenario described, with five evaluators who had different competencies, this could be considered a weakness of the study; however, it is a reality in the daily practice at our services, especially university services.<sup>8,9,18</sup> Another factor that affects the

kappa values is the number of categories to be evaluated, as in the study herein with the classification of fractures and surgical approaches to be selected. In addition, the 20 patients who were studied had different types of tibial plateau fractures (Figure 1). According to the COnsensus-based Standards for the selection of health Measurement INstruments (COSMIN checklist) for reliability studies, the evaluation of 20 patients based on 2D CT scans and 3D printing could be considered to be a small sample and a weakness of the study because a minimum of 50 participants is recommended for reliability studies.<sup>20,21</sup> Furthermore, because the duration of the evaluation time every evaluator took was not standardized, it would be considered another weakness of this study.

The type of instrument used affects reliability, whether using kappa or the intraclass correlation coefficient. In

 Table 2.
 Interobserver Reliability With 2D CT: 3D Printing for the Luo Classification and Surgical Approach

 Compared With the Trauma Surgeon

	CT/Classification			3D/Classification			CT/Approach			3D/Approach		
Evaluator	Kappa	95% CI	%	Kappa	95% CI	%	Kappa	95% CI	%	Kappa	95% CI	%
Knee surgeon	0.24 <sup>a</sup>	0.14-0.36	35	0.30	0.17-0.41	40	0.32	0.21-0.38	45	0.56	0.56-0.67	65
Orthopaedic surgeon	0.25 <sup>a</sup>	0.14-0.43	35	0.44	0.29-0.66.	55	0.24	0.16-0.35	40	0.36	0.27-0.49	50
Third-year resident	0.22 <sup>a</sup>	0.19-0.43	30	0.10	0.03-0.17	20	0.20	0.10-0.31	25	0.27	0.25-0.33	45
Fourth-year resident	0.20 <sup>a</sup>	0.10-0.31	30	0.16	0.09-0.23	25	0.34	0.21-0.46	45	0.63	0.53-0.73	70

2D = two-dimensional, 3D = three-dimensional printing, CI = confidence interval  ${}^{a}P < 0.01$ .

this study, the use of images such as 2D CT and 3D printing resulted in better reproducibility values than radiography.<sup>5,9,22</sup> The use of 3D printing models by this reliability study is a strength because it provides better information and aids in decision-making by decreasing interpretation errors.

Reliability studies of the classification of tibial plateau fractures have used instruments such as plain radiography.<sup>13,23,24</sup> In addition, Schatzker AO/OTA<sup>25</sup> obtained intraobserver kappa values that were between moderate (0.41-0.60) and substantial (0.61-0.80).<sup>19</sup> Studies that use 2D or 3D printing improve intraobserver and interobserver reproducibility,<sup>25-27</sup> with kappa values between slight (0.00-0.20) when evaluating 2D CT scans and moderate (0.41 a 0.60) when evaluating 3D printing models.<sup>19</sup> Furthermore, the kappa value for intraobserver and interobserver reproducibility has been reported to be better with 3D printing.<sup>7</sup>

Normally, the kappa value for interobserver reliability<sup>17</sup> is not high for the different classifications of fractures, ranging from 0.3 (fair) to 0.7 (substantial).<sup>9,27</sup> The interpretation of reliability depends on the context of the evaluator and the instrument used for its evaluation in the clinical practice. In addition, high reliability does not indicate high precision.<sup>8,9</sup>

Innovations in diagnostic imaging have helped with clinical and surgical decisions, which has benefited the treatment of tibial plateau fractures and decreased measurement error in reliability studies.

The 3D printing provides more information, facilitating planning and decision-making for tibial plateau fractures, which are more complex because of the increased energy that they cause. Furthermore, reliability as represented by intraobserver and interobserver reproducibility results in higher and more reproducible kappa using 3D printing coefficient values.

### Conclusion

3D imaging helps to interpret and analyze tibial plateau fractures on the axial, sagittal, and coronal planes by decreasing measurement error.<sup>22</sup> It also improves intraobserver and interobserver reliability in the classification of these fractures and even more improvement is gained in the surgical approach that is selected. The use of 3D printing and its usefulness are helpful to decision-making when providing emergency trauma services to patients with intraarticular fractures such as those of the tibial plateau.

### Acknowledgements

Thank you to the Grupo de investigación de cirugía y especialidades (GRICES) e INTERFAZ (Diseño Industrial), Universidad Industrial de Santander, and to the professors and residents in the orthopedics and traumatology postgraduate program at the Universidad Industrial de Santander.

#### References

1. Elsoe R, Larsen P, Nielsen NPH, Swenne J, Rasmussen S, Ostgaard SE: Population-based epidemiology of tibial plateau fractures. *Orthopedics* 2015;38:e780-e786.

2. Kokkalis ZT, Iliopoulos ID, Pantazis C, Panagiotopoulos E: What's new in the management of complex tibial plateau fractures? *Injury* 2016;47: 1162-1169.

3. Júnior MK, Fogagnolo F, Bitar RC, Freitas RL, Salim R, Jansen Paccola CA: Tibial plateau fractures. *Rev Bras Ortop (Engl Ed)* 2009;44: 468-474.

4. Bishop J, Githens M: Surgical treatment of posterior tibial plateau fractures. *Oper Tech Orthop* 2015;25:242-247.

5. Millar SC, Arnold JB, Thewlis D, Fraysse F, Solomon LB: A systematic literature review of tibial plateau fractures: What classifications are used and how reliable and useful are they? *Injury* 2018;49:473-490.

 Chen H, Chang S, Pan J: Recent progress in the diagnosis and treatment of posterior tibial plateau fractures. *Int J Clin Exp Med* 2015;8: 5640-5648.

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

8. Bland JM, Altman DG: Statistics Notes: Validating scales and indexes. *BMJ* 2002;324:606-607.

9. Audigé L, Bhandari M, Kellam J: How reliable are reliability studies of fracture classifications? A systematic review of their methodologies. *Acta Orthop Scand* 2004;75:184-194.

10. Ardila CC, López CI, Martínez JM, Meléndez GL, Navarro DC, Galeano CF: Study for development of a patient-specific 3D printed craniofacial medical device: Design based on 3D virtual biomodels/CAD/RP. *Proc CIRP* 2018;70:235-240.

11. Huang H, Hsieh MF, Zhang G, et al: Improved accuracy of 3D-printed navigational template during complicated tibial plateau fracture surgery. *Australas Phys Eng Sci Med* 2015;38:109-117.

12. Faria C, Pinho D, Santos J, Gonçalves L, Lima R: Low cost 3D printed biomodels for biofluid mechanics applications. *J Mech Eng Biomech* 2018; 3:1-7.

13. Luo CF, Sun H, Zhang B, Zeng BF: Three-column fixation for complex tibial plateau fractures. *J Orthop Trauma* 2010;24:683-692.

14. Ulbrich C, Zavaglia C, Leivas T, Teixeira F. Medical application of rapid prototyping in orthopedics surgical planning, in Bártolo P, ed: *Innovative Developments in Virtual and Physical Prototyping*. London, UK, Taylor and Francis Group, 2012, pp 85-87.

15. López CI, Bautista LE, Mantilla C: A novel approach to learning virtual engineering in the medical devices process. *Educ Inform Technol* 2020;25: 5703-5729.

16. Martinez J, López C, Garnica I, Murillo A: Decision engineering, in Stark J, ed: *Product Lifecycle Management (Volume 4): The Case Studies*. Cham, Switzerland, Springer International Publishing, 2019, pp 282.

#### Jose Osma-Rueda, MD, MSc, et al

**Research** Article

18. Sim J, Wright CC: The kappa statistic in reliability studies: Use, interpretation, and sample size requirements. *Phys Ther* 2005;85:257-268.

19. Landis JR, Koch GG: The measurement of observer agreement for categorical data. *Biometrics* 1977;33:159.

20. Mokkink LB, Terwee CB, Patrick DL, et al: The COSMIN study reached international consensus on taxonomy, terminology, and definitions of measurement properties for health-related patient-reported outcomes. *J Clin Epidemiol* 2010;63:737-745.

21. Müller ME, Koch P, Nazarian S, Schatzker J: Tibia/Fibula = 4, in *The Comprehensive Classification of Fractures of Long Bones*. Berlin, Heidelberg, Springer-Verlag, 1990, pp148-191.

22. Karanicolas PJ, Bhandari M, Kreder H, et al: Evaluating agreement: Conducting a reliability study. *J Bone Joint Surg Am* 2009;91(suppl 3):99-106. 23. Hohl M, Luck JV: Fractures of the tibial condyle; a clinical and experimental study. *J Bone Joint Surg Am* 1956;38:1001-1018.

24. Duparc J, Ficat P: Articular fractures of the upper end of the tibia [in French]. *Rev Chir Orthop Reparatrice Appar Mot* 1960;46: 399-486.

25. Gicquel T, Najihi N, Vendeuvre T, Teyssedou S, Gayet LE, Huten D: Tibial plateau fractures: Reproducibility of three classifications (Schatzker, AO, Duparc) and a revised Duparc classification. *Orthop Traumatol Surg Res* 2013;99:805-816.

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

27. Martínez-Rondanelli A, Escobar-González SS, Henao-Alzate A, Martínez-Cano JP: Reliability of a four-column classification for tibial plateau fractures. *Int Orthop* 2017;41:1881-1886.