The Effect of Three-Dimensional Printed Models on Reliability of the Assessment of the Lauge-Hansen Classification System

Christopher J. Berkelbach, Stephanie L. Golding, Emily Lobos, Kalen Farr, Shelby Hatzinger, Jacqueline Higgins, Samuel Straus

Student Members, ACFAS Club, Temple University School of Podiatric Medicine, Philadelphia, Pennsylvania

Literature Review and Statement of Purpose

Classification systems for orthopedic injuries have been developed to assist physicians in planning their management and predicting outcomes [1]. It is important that commonly used classification systems are both reliable and valid. There is a growing concern that fracture classification systems should be formally validated before they are recommended for use in practice.

Ankle injuries are one of the leading causes of musculoskeletal emergency department visits in the U.S. [2]. Roughly 95% of all ankle fractures can be grouped according to category and stage under the Lauge-Hansen (LH) classification system [3], which was developed by Danish radiologist Niels Lauge-Hansen in 1942 based on findings from experimentally produced fractures on cadaveric models [4]. The LH classification system is thought of by many foot and ankle surgeons as an important tool for the understanding and treatment of ankle fractures. However, many issues have been raised concerning its reliability and reproducibility over the years [3, 5-8]. At this point in time, the flexibility for correction to the existing LH classification system is limited. Attempts have been made, however, to study external effects on the validity and reliability of the LH system. Rasmussen et al. observed the benefits of instruction on increasing agreement using the LH system in classifying ankle fracture radiographs [9]. Rodriguez et al., however, reported a low level of accuracy in the use of the LH system when radiographs were matched to corresponding video clips of the actual injury [10]. Hasenstein et al. demonstrated high levels of satisfaction with the incorporation of 3D printed models of ankle fractures into podiatric medical education [11]. Misselyn et al. has recently reported improved interobserver reliability of the Sanders classification with the addition of 3D printed models of calcaneal fractures [12]. It is with this evidence that the current project was conceived.

The purpose of this study was to evaluate the effect on interobserver and intraobserver reliability in the LH classification system with the addition of 3D printed models. The primary objective of the study was to compare the interobserver and intraobserver reliability of the LH classification system for x-ray films and 3D printed models. The secondary objectives were to compare the strength of the effect of 3D printed models on reliability between attending physicians, resident physicians and podiatry students, and to identify patterns of over- or under-staging of 3D prints compared to x-ray films.

Methodology

The 3D printed models were utilized from a previous IRB approved protocol evaluating their use as an educational instrument [11]. Models were approximately 20cm high and could be durably handled. A total of thirty-four participants were recruited to take part in the study: five ABFAS certified foot and ankle surgeons, seventeen resident physicians, and twelve fourth year podiatric medical students. All participants reported feeling comfortable classifying ankle fractures using the LH system. Participants were given the opportunity to review the classification system prior to their evaluations as referenced from Tartaglione et al. [8]. Participants were also given the opportunity to examine a 3D printed model of a pilon fracture to aid their familiarity with 3D printed models. Participants evaluated a series of 18 items alternating between 3D printed models and x-ray films of 7 unique ankle fractures classifiable by the LH system. Of the 7 ankle fractures, 4 were classified as SER IV (Items A-D), 2 were classified as PER IV (Items E-F), and 1 classified as PAB 1 (Item G) by the first author (C.B.). Multiple x-ray views were provided for each fracture to allow determination of the full extent of the injury. Two 3D printed models and x-ray films were presented twice in order to assess intrarater reliability. The order of presentation was randomized, but the same for each participant. Participants were asked to assign a LH category and stage for each fracture by marking the corresponding box on an answer sheet. No time limit was imposed.

The distribution of LH classification scores were then compared between the x-ray films and 3D prints. Two quantifications of agreement are reported. First, we recorded the percentage agreement, which measured the rate as which participants gave the same score. Second, we recorded Fleiss's fixed-marginal multirater kappa, which is a measure of the degree of interrater agreement beyond the level of chance. The Landis and Koch benchmark scale was used to interpret the strength of agreement for Fleiss kappa values, according to the following: 0 = poor agreement; 0 to 0.2 – slight agreement; 0.21 to 0.4 – fair agreement; 0.41 to 0.6 – moderate agreement; 0.61 to 0.8 – substantial agreement; 0.81 to 1.0 – almost perfect agreement [13]. The analyses were performed using a free Online Kappa Calculator available at http://justus.randolph.name/kappa. [14]. Representative examples of different ankle fractures as 3D prints are provided in Figure 1.

Results

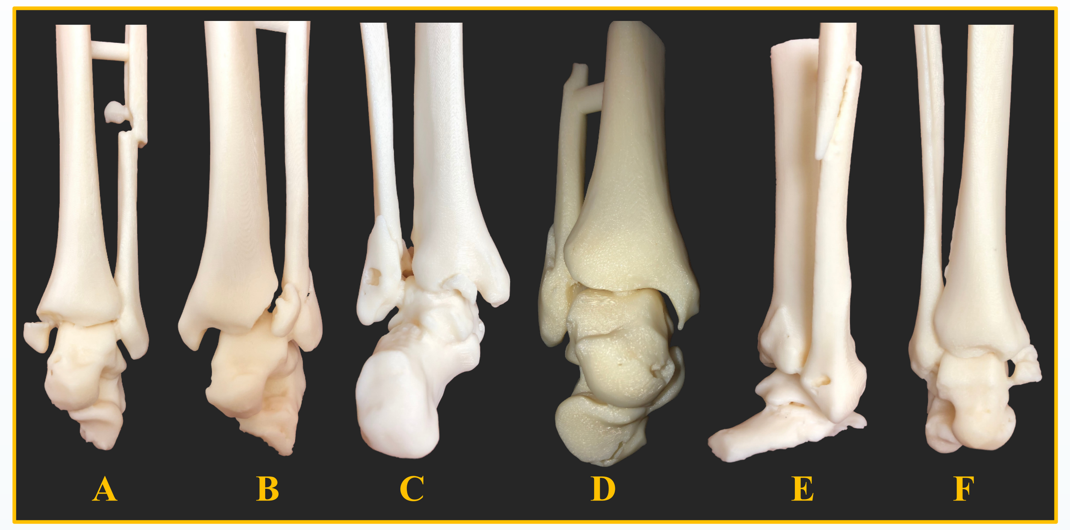

Figure 1. Examples of 3-dimensional prints of ankle fractures. **A)** PER injury with midshaft fibular fracture and medial malleolar avulsion fracture. **B)** SER injury with spiral oblique fibular fracture and Tillaux fracture of the anterior lateral margin of the distal tibia. **C)** SER injury with posterior malleolar fracture. **D)** SER injury with spiral oblique fracture of the fibula and widening of the medial clear space with fleck avulsion fracture of the medial malleolus **E)** PER injury with posterior malleolar fracture and midshaft fibular fracture. **F)** PAB injury with transverse medial malleolar fracture.

	Agreement to Reference Standard	
	X-Rays	3D Prints
All observers (N = 34)	70.2%	61.7%
Attending physicians (N = 5)	65.7%	60.0%
Residents (N = 17)	74.8%	65.5%
Students (N = 12)	70.2%	59.5%

Table 1. With the x-ray images, the percent agreement with the reference standard was higher than that with the 3D prints across all groups, with the most significant difference seen in students followed by residents.

	Interobserver agreement	
	X-Rays	3D Prints
All observers (N = 34)		
Agreement	54.2%	56.0%
Fleiss kappa (95% CI)	0.37 (0.31-0.43)	0.43 (0.38-0.49)
Attending physicians (N = 5)		
Agreement	51.4%	57.1%
Fleiss kappa (95% CI)	0.30 (0.10-0.49)	0.43 (0.26-0.60)
Residents (N = 17)		
Agreement	54.0%	58.7%
Fleiss kappa (95% CI)	0.35 (0.26-0.45)	0.46 (0.38-0.54)
Students (N = 12)		
Agreement	53.7%	54.6%
Fleiss kappa (95% CI)	0.39 (0.29-0.48)	0.42 (0.32-0.51)

Table 2. When participants were asked to categorize the ankle fractures according to LH classification with class and stage using the 3D printed models, **there was an increased interobserver agreement compared with the x-ray images. The Fleiss kappa value was also increased.** The improvement in interobserver agreement and Fleiss kappa with the 3D prints over x-ray films was seen across all groups and was most marked among attending physicians. **All Fleiss kappa values for interobserver agreement for x-rays fell between 0.21 to 0.4, interpreted as fair agreement. While all Fleiss kappa values for interobserver agreement for 3D prints fell between 0.41 to 0.6, interpreted as moderate agreement.**

	% Perfect Agreement	
	X-Rays	3D Prints
All observers (N = 34)	64.7% (44/68)	77.9% (53/68)
Attending physicians (N = 5)	90.0% (9/10)	90.0% (9/10)
Residents (N = 17)	52.9% (18/34)	82.4% (28/34)
Students (N = 12)	70.8% (17/24)	66.7% (16/24)

Table 3. When observers classified identical (repeat) ankle fractures, the result was an **increased intraobserver perfect agreement with the 3D prints compared to the x-ray films.** This effect was most marked among resident physicians, however was not seen across all groups.

	All observers (N = 34)	Attendings (N = 5)	Residents (N = 17)	Students (N = 12)
Total fractures classified	238 (100%)	35 (100%)	119 (100%)	84 (100%)
Print and film assigned different category	46 (19.3%)	8 (22.9%)	24 (20.2%)	14 (16.7%)
Print and film assigned same category	192 (80.7%)	27 (77.1%)	95 (79.8%)	70 (83.3%)
Print and film assigned same category	192 (100%)	27 (100%)	95 (100%)	70 (100%)
and same stage	122 (63.5%)	18 (66.7%)	63 (66.3%)	41 (58.6%)
and different stage	70 (36.5%)	9 (33.3%)	32 (33.7%)	29 (41.4%)
Print and film assigned same category and different stage	70 (100%)	9 (100%)	32 (100%)	29 (100%)
Print higher stage than film	11 (15.7%)	1 (11.1%)	4 (12.5%)	6 (20.7%)
Print lower stage than film	59 (84.3%)	8 (88.9%)	28 (87.5%)	23 (79.3%)

Table 4. Across all participant groups, the **same category** was assigned for both the 3D print and x-ray film of the same fracture **80.7%** of the time (192/238), while the **same category and stage** were assigned for both the 3D print and x-ray film of the same fracture **51.3%** of the time (122/238). **Of those fractures assigned the same category but different stage, a lower stage was assigned for the 3D print 84.3% of the time overall.** This strength of this effect correlated with level of experience and expertise.

	Item A	Item B	Item C	Item D	Item E	Item F	Item G
All observers (N = 34)							
Agreement for X-ray	0.583	0.777	0.679	0.237	0.681	0.681	0.159
Agreement for 3D print	0.781	0.453	0.597	0.884	0.681	0.376	0.148
Attending physicians (N = 5)							
Agreement for X-ray	0.600	0.600	0.600	0.200	0.600	1.000	0.000
Agreement for 3D print	1.000	0.400	0.200	1.000	1.000	0.400	0.000
Residents (N = 17)							
Agreement for X-ray	0.669	0.772	0.676	0.309	0.581	0.581	0.191
Agreement for 3D print	0.882	0.493	0.691	0.882	0.669	0.316	0.176
Students (N = 12)							
Agreement for X-ray	0.439	0.833	0.682	0.121	0.833	0.682	0.167
Agreement for 3D print	0.561	0.364	0.682	0.833	0.561	0.545	0.273

Table 5. Agreement was found to be superior for 3D print over x-ray film across all groups for certain fractures (Items A & D), while agreement was superior for x-ray film over 3D print for other fractures (Items B & F). Overall, agreement for classification of SER fractures (Items A-D) was superior for 3D prints compared to x-ray films (0.679 versus 0.569) while agreement for classification of PER fractures (Items E & F) was superior for x-ray films compared to 3D prints (0.507 versus 0.402).

Discussion

The objective of this investigation was to evaluate the effect of 3D printed models on the reliability of the Lauge-Hansen classification system.

Interobserver reliability was found to fall within the range of **"moderate agreement" for 3D prints, which is improved compared to "fair agreement" for x-ray films**. We find this to be an important finding with respect to preoperative planning of high stage ankle fractures, which often require open reduction and internal fixation (ORIF). The best results for ORIF require good preoperative assessment of the fracture. High stage ankle fractures are usually complex with comminution, displacement and/or rotational abnormalities. The primary aim in treating ankle fractures is to restore the morphology and integrity of the ankle mortise. To achieve this, good reduction and fixation of ankle fractures are essential, but sometimes difficult to perform. X-ray along with 3D CT imaging are the current gold standards for preoperative planning of high stage ankle fractures. **However, the advent of 3D printing potentially opens the door for even better preoperative assessment of these fractures.** Holding fracture models in their hands would allow surgeons to get a clear determination of the personality of the fracture without having to imagine the pathology in their minds based on images seen on a computer screen. Studies have shown that 3D printed models can improve understanding of fracture pathology, leading to reduced operative times, increased surgical efficiency, and reduced iatrogenic complications [15-17]. Furthermore, Sha et al. demonstrated the use of 3D templates to improve outcomes for lateral ankle ligament reconstruction surgery [18]. Chung et al. described the use of 3D templates for selection of plates and screws, as well as for guiding reduction and screw trajectories for distal tibial fractures [19].

A limitation of relying on 3D printed models for preoperative assessment is the resolution and accuracy of the printed model itself. **Our study did not, however, reveal decreased accuracy in classifying the 3D prints when compared to 2D x-ray, demonstrated in Table 3 by increased percent perfect agreement with 3D prints**. Another limitation of the 3D printed models is simply lack of familiarity and experience with handling them. This may explain why participants' responses were more likely to agree with the reference standard for x-ray versus 3D print (Table 1). It may also explain why **3D prints were assigned a lower stage 84.3% of the time** when compared with classification assigned for an x-ray of the same fracture (Table 4). This lack of experience with 3D printed models may have caused some participants to, for example, neglect an injury to the posterior inferior tibiofibular ligament, thus classifying a PER IV injury as a PER III.

In conclusion, the results of this investigation provide new evidence on the utility of 3D printed models of ankle fractures for improved reliability of the Lauge-Hansen classification system. Studies such as ours may continue to open the door for further study on 3D printed models as preoperative tools to improve ankle fracture assessment and classification.

References

[1] Audigé L, Bhandari M, Kellam J. How reliable are reliability studies of fracture classifications? A systematic review of their methodologies. Acta Orthopaedica Scandinavica. 2004 Jan 1;75(2):184-94.

[2] Shibuya N, Davis ML, Jupiter DC. Epidemiology of foot and ankle fractures in the United States: an analysis of the National Trauma Data Bank (2007 to 2011). The Journal of Foot and Ankle Surgery. 2014 Sep 1;53(5):606-8.

[3] Shariff SS, Nathwani DK. Lauge-Hansen classification—a literature review. Injury. 2006 Sep 1;37(9):888-90.

[4] Yde J. The Lauge Hansen classification of malleolar fractures. Acta Orthopaedica Scandinavica. 1980 Jan 1;51(1-6):181-92.

[5] Chen YM, Fei YX, Ming MJ, Xia Y, Wang T, li XX, Jie YY, Hai XJ, Ye J, Yi TZ, Qi FY. Evaluating the reliability and reproducibility of the AO and Lauge-Hansen classification Systems for Ankle Injuries. Orthopedics. 2015 Jul 10;38(7):e626-30.

[6] Nielsen JØ, Dons-Jensen H, Sørensen HT. Lauge-Hansen classification of malleolar fractures: an assessment of the reproducibility in 118 cases. Acta orthopaedica Scandinavica. 1990 Jan 1;61(5):385-7.

[7] Thomsen NO, Overgaard S, Olsen LH, Hansen H, Nielsen ST. Observer variation in the radiographic classification of ankle fractures. The Journal of bone and joint surgery. British volume. 1991 Jul;73(4):676-8.

[8] Tartaglione JP, Rosenbaum AJ, Abousayed M, DiPreta JA. Classifications in brief: Lauge-Hansen classification of ankle fractures.

[9] Rasmussen S, Madsen PV, Bennicke K. Observer variation in the Lauge-Hansen classification of ankle fractures: precision improved by instruction. Acta orthopaedica Scandinavica. 1993 Jan 1;64(6):693-4.

[10] Rodriguez EK, Kwon JY, Chacko AT, Kadzielski JJ, Herder L, Appleton PT. An Update on Assessing the Validity of the Lauge Hansen Classi□ication System for In-vivo Ankle Fractures Using YouTube videos of Accidentally Sustained Ankle Fractures as a Tool for the Dynamic Assessment of Injury. The Harvard Orthopaedic Journal. 2012 Dec; 14: 40-43.

[11] Hasenstein T, Patel KP, Meyr AJ. Printed Three-Dimensional Computerized Tomography Scanned Ankle Fractures as an Educational Instrument. Poster presented at: 2018 American College of Foot and Ankle Surgeons Scientific Conference. 2018 Mar 22-25; Nashville, TN.

[12] Misselyn D, Nijs S, Fieuws S, Shaheen E, Schepers T. Improved Interobserver Reliability of the Sanders Classification in Calcaneal Fractures Using Segmented Three-Dimensional Prints. The Journal of Foot and Ankle Surgery. 2018 May 1;57(3):440-4.

[13] Landis JR, Koch GG. The measurement of observer agreement for categorical data. biometrics. 1977 Mar 1:159-74.

[14] Randolph JJ. Online Kappa Calculator [Computer software]. 2008; Retrieved from http://justus.randolph.name/kappa.

[15] Sanghavi PS, Jankharia BG. Holding versus seeing pathology. Three-dimensional printing of the bony pelvis for preoperative planning of a complex pelvis fracture: A case report. The Indian journal of radiology & imaging. 2016 Jul;26(3):397.

[16] Starosolski ZA, Kan JH, Rosenfeld SD, Krishnamurthy R, Annapragada A. Application of 3-D printing (rapid prototyping) for creating physical models of pediatric orthopedic disorders. Pediatric radiology. 2014 Feb 1;44(2):216-21.

[17] Lal H, Patralekh MK. 3D printing and its applications in orthopaedic trauma: A technological marvel. Journal of clinical orthopaedics and trauma. 2018 Aug 3.

[18] Sha Y, Wang H, Ding J, et al. A novel patient-specific navigational template for anatomical reconstruction of the lateral ankle ligaments. Int Orthop. 2016 Jan 1;40(1):59e64.

[19] Chung KJ, Huang B, Choi CH, Park YW, Kim HN. Utility of 3D printing for complex distal tibial fractures and malleolar avulsion fractures: technical tip. Foot & ankle international. 2015 Dec;36(12):1504-10.

We would like to acknowledge our faculty advisors (Andrew J. Meyr, DPM FACFAS and Laura Sansosti, DPM AACFAS) for their advisory roles with this project and presentation.