# Comparison of the Accuracy of X-ray, 2D-CT, 3D-CT, and Physical Modeling in Classification of Fractures about the Elbow Needing Operative Treatment

**Thomas DeCoster MD[1], Deana Mercer MD[1], and Evan Baldwin MD[2]**

*1. UNM Department of Orthopaedics & Rehabilitation 2. UNM Department of General Surgery*

**Abstract**

The goal of this study is to compare the accuracy of using X-ray, 2D-CT, 3D-CT, and physical modeling in classification of fractures about the elbow as a means of evaluating their relative utility in preoperative workup and treatment planning of fractures.

Ten patients with fractures about the elbow that required operative fixation underwent preoperative X-ray, 2D-CT, 3D-CT, and physical modeling of their injury. Ten orthopaedic physicians classified each injury using each of those four modalities. The answers given by the 10 physicians were compared to an established correct classification for each case, and that data was used to compare the relative accuracy of each modality.

The average accuracy for the given modalities was 62% for X-ray, 76% for 2D-CT, 80% for 3D-CT, and 88% for physical modeling. ANOVA analysis across all modalities revealed findings are statistically significant; however, when compared side by side, only moving from X-ray to 2D-CT yielded significant results.

There was greater percentage correct classification achieved using the more advanced modalities, which therefore may theoretically result in more accurate preoperative planning. However, one must view this finding within the context and limits of this study, which is restricted by the relatively small sample size. Future study into methods of fracture characterization should be done to further evaluate findings such as these, with the goal of promoting better patient outcomes.

**Introduction**

Amongst all fractures in humans, those occurring about the elbow can be quite complex and challenging to treat.[1] Therefore, accurate preoperative radiological characterization of the fracture is important and facilitates the planning and execution of injury management. Prior studies into the value of such preoperative investigations have demonstrated improved injury characterization with three-dimensional (3D-CT) compared to two-dimensional computed tomography (2D-CT) images and radiographs.[1-8] In addition, over the last 20 years there has been significant investigation into the utility of three-dimensional (3D) physical models that are constructed from CT images of bony injuries. These physical models can actually be held in the hand of the physician and may facilitate superior evaluation of fracture characteristics in surgical planning.[7] However, these prior studies have been based upon retrospective data, and the accuracy of the images in particular relied strongly upon recollection and operative notes. Additionally, there have been no studies published that comparatively evaluate the utility of all four modalities (X-ray, 2D-CT, 3D-CT, and physical modeling) in fracture evaluation. Therefore, using fractures about the elbow as the chosen injury type, we will evaluate those modalities by comparing the accuracy of classification of those fractures by orthopaedic surgery physicians using each modality to evaluate a set of cases for which all evaluative methods were obtained. As a study like this has yet to be published, our hope is that through a manageable sample size here at University of New Mexico Hospital (UNMH) we can carry out a successful pilot prompting future larger studies.

**Methods**

In order to obtain a set of cases suitable for this study, we identified 10 adult (18 years of age or older) patients, regardless of sex, race, or ethnicity, as they presented to UNMH with a fracture about the elbow that required operative treatment and underwent both plain film and 2D-CT studies, per standard of care in this case, in the preoperative evaluation of their injury. All of these patients were consented for their involvement in this study and signed an informed consent agreement. We then contacted the UNMH Radiology Department and had the 2D-CT data for each of the 10 cases reconstructed into a 3D-CT representation. Additionally, the 2D-CT data was deidentified and sent via secure connection to Medical Modeling LLC of Golden, Colorado for manufacturing of physical models of the injuries via laser stereolithography. These models were then sent back to the Department of Orthopaedics and Rehabilitation at UNMH. The data modeling company then destroyed the data used to manufacture the model. Concurrently, the patient underwent surgery at UNMH. At this point, we had created a set of 10 cases for which all 4 imaging modalities were obtained.

To evaluate the utility of the 4 imaging modalities, a computer program was written by Evan Baldwin (EB) using Microsoft Access that allows a user to answer sequential questions about a given injury and proceed through the Orthopaedic Trauma Association (OTA) fracture classification scheme to arrive at a single fracture classification value. The program then transferred all of the selections and the final answer into a database. A total of 10 orthopaedic surgery physicians at UNMH, not directly involved in the original patient case, were then asked to use the computer program to classify the injuries in the 10 cases. These 10 physicians went through all 10 cases, classifying the injury using each of the 4 modalities, thus creating 10 points of data for each of the modalities that could be used to compare the accuracy of the imaging modality. To avoid any confusion using the OTA classification, each user was given a copy of the classification scheme, complete with illustrations, as published in the *Journal of Orthopaedic Trauma* (Vol. 21, Number 10 Supplement, Nov/Dec 2007). EB was present with all of the physicians involved to ensure the data was collected accurately.

In order to process the data, it was necessary to develop a gold standard for correct classification of the fracture about the elbow in each of the 10 cases against which the data for that case could be compared to develop the relative accuracy of a given modality. This was done by having an attending physician of upper extremity specialization classify the fracture using all 4 modalities to create a single answer and comparison to operative notes for the given case.

**Results**

The answers entered into the computer by the 10 physicians going through the cases were compared to the gold standard in order to determine the level of accuracy that was obtained using that modality. The results are presented in Table 1.

Overall, the average accuracy for the given modalities was 62% for X-ray, 76% for 2D-CT, 80% for 3D-CT, and 88% for physical modeling. Graphical representation of the progression of increasing correct classifications can be seen in Figure 1. ANOVA (analysis of variance) testing ($\alpha = 0.05$) across all the modalities revealed a p value of very much less than 0.05 (0.0003).

However, when comparisons from one level of evaluation to the next are made, the significance is notably different: X-ray/2D-CT p=0.026, 2D-CT/3D-CT p=0.433, and 3D-CT/modeling p=0.136.

**Discussion**

En masse, the relative percentage correct achieved using the more advanced modalities to classify the fractures was greater, and therefore, theoretically more likely to result in more accurate preoperative planning. However, one must view this finding within the context

Table 1.

*Percent Correct Fracture Classification by Case and Modality*

| | | Percent correct classification by evaluative modality | | | |
|---|---|---|---|---|---|
| Case Number | Correct OTA Classification | X-ray | 2D-CT | 3D-CT | Physical model |
| 1 | 21-B1.1 (1) | 70% | 80% | 90% | 90% |
| 2 | 13-B1.1 (1) | 50% | 70% | 80% | 80% |
| 3 | 21-B2.1 (2) | 70% | 70% | 80% | 80% |
| 4 | 21-B1.1 (4) | 70% | 80% | 80% | 100% |
| 5 | 21-B1.3 (3) | 50% | 60% | 80% | 90% |
| 6 | 21-C2.3 | 70% | 80% | 80% | 80% |
| 7 | 21-C1.2 | 40% | 60% | 90% | 100% |
| 8 | 21-C1.2 | 80% | 90% | 60% | 60% |
| 9 | 21-C2.1 | 50% | 70% | 70% | 100% |
| 10 | 21-B2.1 (2) | 70% | 100% | 90% | 100% |

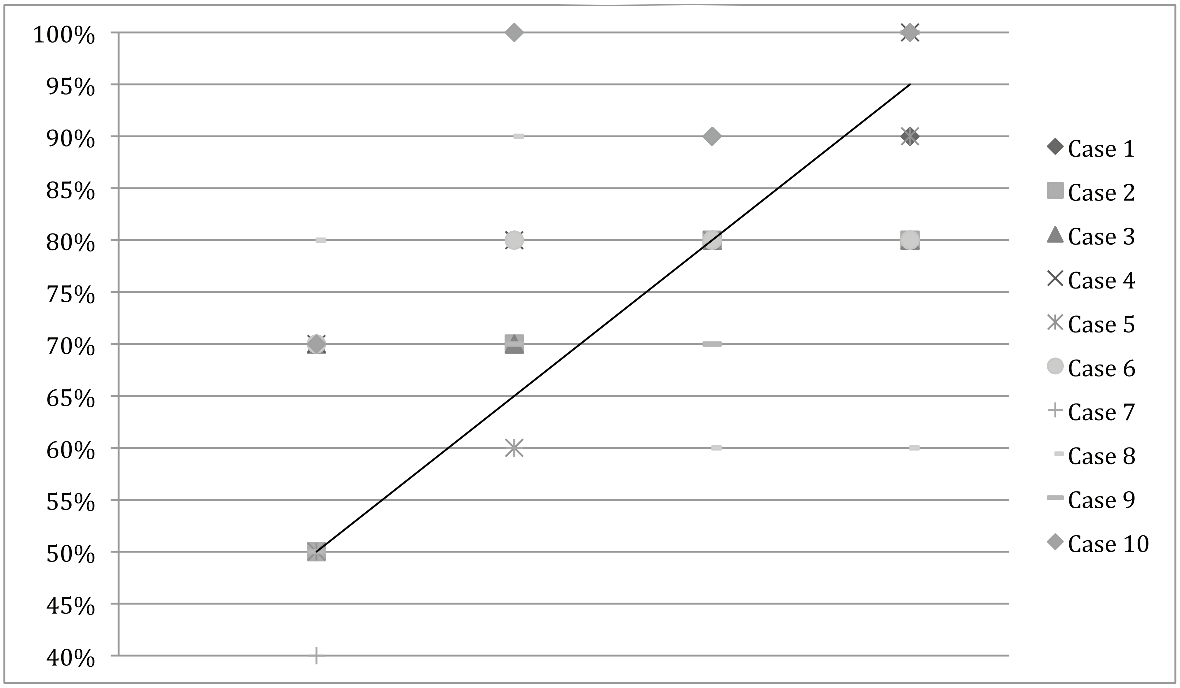

*Figure 1*: Percent correct fracture classification, proceeding from X-ray to 2D-CT to 3D-CT to Physical Modeling, with average trend line.

and limits of this study. First, we must acknowledge that the relative superiority is seen only using ANOVA across all of the modalities, and that when viewed in a post hoc manner, only moving from X-ray to 2D-CT produces a statistically significant finding. This is an important distinction because in cases of fracture about the elbow that require operative treatment, 2D-CT is already standard of care in most major medical centers with access to such imaging. This finding sheds light on debate regarding pursuit of more advanced imaging once standard of care is achieved – if a surgeon is not better able to classify a fracture, and therefore theoretically better carry out surgical planning, is it worthwhile to incur larger cost to the healthcare system to obtain information that does not significantly add value? While this may seem rhetorical, answering such a question should be done within not only the setting of today's healthcare infrastructure, but also that of the future, where the cost of pursuing more advanced imaging modalities might not add significant fiscal burden. In such an instance, expanding standard of care to include 3D-CT and/or physical modeling of bony injuries could be enacted as a means of ensuring every effort to promote patient well-being and safety is undertaken.

An interesting finding, although not easily addressed statistically, is the decreased correct percent classification in 2 of the cases that had a nondisplaced fracture fragment. Correct classification was more, or equally often, achieved with X-ray and 2D-CT as compared to 3D-CT and physical modeling in these instances. This conceivably occurred due to the inability of those 2 more advanced modalities to communicate fracture of bone without disruption of the natural contours, and the comparatively less discrete and sensitive manner in which the data is presented.

Furthermore, the findings of this study must be considered against its limitations. This study would be much more powerful if it not only had more patient cases, but also had many more physicians participating in the classification. This could be addressed in future studies where a multicenter approach might be better suited to attain large numbers. Every effort was made to eliminate all reasonable bias within the study, but it is possible that some may have occurred. We recognize that it is possible a participating physician could have unknowingly classified a fracture that he or she had previously been involved with, thus skewing the response. Although we attempted to control for physician inexperience with the OTA classification scheme by creating a user friendly computer program and providing supplemental visual materials, we did not control for experience with upper extremity trauma or naiveté with the system. Lastly, it is possible that the X-ray and CT data may not have been of identical fracture patterns, as the time interval between those imaging sessions and consequent patient movement could have disrupted the location of bones and fragments in a given case.

In sum, evaluation of our current methods of fracture assessment should be carried out on an ongoing basis, as should comparative study of our current standards of practice against new and emerging technologies and ideas. This study, while small in size, demonstrates that there is inequity in the information that practitioners receive from different imaging modalities when characterizing a bony injury. All methods-X-ray, 2D-CT, 3D-CT, and physical modeling-have advantages and drawbacks that should be further assessed in future study. Specifically, research is needed to investigate the generalizability of these findings to fractures outside of the elbow, relative costs to the healthcare system incurred when advanced modalities are employed, and whether or not better classifying a fracture alters treatment planning or patient outcomes.

**References**


1. Castagno AA, Shuman WP, Kilcoyne RF, Haynor DR, Morris ME, Matsen FA. Complex fractures of the proximal humerus: role of CT in treatment. *Radiology.* December 1987;165(3):759-762.

2. Siebenrock KA, Gerber C. The reproducibility of classification of fractures of the proximal end of the humerus. *J Bone Joint Surg Am*. December 1993;75(12):1751-1755.

3. Kuhlman JE, Fishman EK, Ney DR, Magid D. Complex shoulder trauma: three-dimensional CT imaging. *Orthopedics*. November 1988;11(11):1561-1563.

4. Kilcoyne RF, Shuman WP, Matsen FA III, Morris M, Rockwood CA. The Neer classification of displaced proximal humeral fractures: spectrum of findings on plain radiographs and CT scans. *Am J Roentgenol*. May 1990;154(5):1029-1033.

5. Smith RM. The classification of fractures. *J Bone Joint Surg Br.* July 2000;82(5):625-626.

6. Katz MA, Beredjiklian PK, Bozentka DJ, Steinberg DR. Computed tomography scanning of intra-articular distal radius fractures: does it influence treatment? *J Hand Surg [Am]*. May 2001;26(3):415-421.

7. Edelson G, Kelly I, Vigder F, Reis ND. A three-dimensional classification for fractures of the proximal humerus. *J Bone Joint Surg [Br]*. April 2004;86(3):413-425.

8. Borrelli J Jr, Goldfarb C, Catalano L, Evanoff BA. Assessment of articular fragment displacement in acetabular fractures: a comparison of computerized tomography and plain radiographs. *J Orthop Trauma*. August 2002;16(7):449,456; discussion 456-457.

9. Cole RJ, Bindra RR, Evanoff BA, Gilula LA, Yamaguchi K, Gelberman RH. Radiographic evaluation of osseous displacement following intra-articular fractures of the distal radius: reliability of plain radiography versus computed tomography. *J Hand Surg [Am]*. September 1997;22(5):792-800.

10. Marsh JL, Slongo TF, Agel J, et al. Orthopaedic Trauma Association Classification, Data and Outcomes Committee. Fracture and dislocation classification compendium-2007. *J Orthop Traum.* November/December 2007;21(10):S1-S6.

11. Andersen DJ, Blair WF, Steyers CM Jr, Adams BD, el-Khouri GY, Brandser EA. Classification of distal radius fractures: an analysis of interobserver reliability and intraobserver reproducibility. *J Hand Surg [Am].* July1996;21(4):574-582.

12. Bernstein J, Adler LM, Blank JE, Dalsey RM, Williams GR, Iannotti JP. Evaluation of the Neer system of classification of proximal humeral fractures with computerized tomographic scans and plain radiographs. *J Bone Joint Surg Am.* September 1996;78(9):1371-1375.

13. Horn BD, Rettig ME. Interobserver reliability in the Gustilo and Anderson classification of open fractures. *J Orthop Trauma*.1993;7(4):357-360.