Distal humerus traction radiographs: is the inter-observer and intra-observer reliability comparable with computed tomography

Joseph D. Galloway M.D., Stephen J. Shymon M.D., Mark R. Adams M.D., Mark C. Reilly M.D., Michael S. Sirkin M.D., Jeremy Hreha M.D., Michael T. Jung M.D., Naji Madi M.D., Brianna L. Siracuse M.D., Irfan Ahmed M.D., Michael M. Vosbikian M.D.

Rutgers New Jersey Medical School; Department of Orthopaedics

Introduction

- Distal humerus fracture classification and characterization are difficult on plain radiographs
- Traction radiographs and CT have been proposed to improve characterization.
- Traction radiographs have been shown to be comparable to 2D-CT in classification and characterization in distal radius fractures
- Reliable characterization and classification is needed for planning and research.

Methods

- 18 Distal humerus fracture plain films, traction films, and 2D-CT were randomized into multiple Presentations.
- 11 orthopaedic surgeons of varied training levels reviewed all 18 patients on 4 separate occasions and classified fractures by 2 major schemes and answered 6 discrete questions about fracture characteristics and treatment approach
- Interobserver and intraobserver reliability was calculated by kappa statistic

Results

- The interobserver agreement of classification of distal humerus fractures by Traction or 2DCT is fair to moderate, and not significantly improved over plain radiographs.
- The intraobserver agreement for classification however is substantial for both AO and Jupiter-Mehne by either traction or 2D-CT

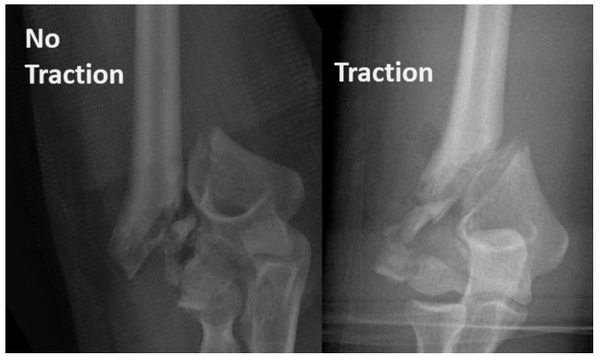
Overall Cohort	**Interobserver Reliability***			**Intraobserver Reliability***		
Question	Plain + Traction	Plain + 2D-CT	p-value	Plain + Traction	Plain + 2D-CT	p-value
AO Classification	**0.30 (0.25-0.34)**	**0.42 (0.39-0.46) †**	**< 0.01**	0.72 (0.67-0.77)	0.70 (0.65-0.76)	0.63
Jupiter Mehne Classification	0.32 (0.28-0.35)	0.30 (0.26-0.33)	0.31	0.73 (0.70-0.77)	0.75 (0.72-0.78)	0.43
Coronal Fracture	**0.20 (0.13-0.26)**	**0.34 (0.27-0.41) †**	**< 0.01**	0.41 (0.32-0.51)	0.57 (0.41-0.73)	0.09
Articular Comminution	0.50 (0.44-0.56)	0.56 (0.55-0.64)	0.05	0.78 (0.68-0.88)	0.64 (0.57-0.72)	0.07
Metaphyseal Comminution	0.21 (0.15-0.28)	0.20 (0.14-0.27)	0.74	0.52 (0.44-0.61)	0.48 (0.34-0.62)	0.68
Separate Articular Fragment	**0.24 (0.17-0.31)**	**0.36 (0.30-0.42) †**	**< 0.01**	0.60 (0.43-0.77)	0.57 (0.42-0.72)	0.78
Impacted Stable Articular Fragment	0.19 (0.09-0.29)	0.15 (0.08-0.22)	0.51	**0.67 (0.48-0.85)**	**0.40 (0.26-0.54)**	**0.01**
Do you need an Olecranon Osteotomy	0.29 (0.26-0.33)	0.38 (0.31-0.44)	0.03	0.51 (0.36-0.67)	0.70 (0.54-0.86)	0.11
Decide from 6 treatment options	**0.07 (0.02-0.11)**	**0.00 (-0.04 - 0.04)**	**0.02**	0.60 (0.45-0.76)	0.63 (0.50-0.76)	0.69

*The values are given as the kappa coefficient, with the 95% confidence interval. †Significant increase in agreement.

Subgroup Analysis

- In trainees, 2D-CT improved the interobserver reliability of:
  - Detection of coronal fracture lines from slight to moderate ($\kappa_{Traction}$ = 0.16; $\kappa_{2D\text{-}CT}$ = 0.48)
  - Detection of articular comminution from fair to substantial ($\kappa_{Traction}$ = 0.39; $\kappa_{2D\text{-}CT}$ = 0.66)
  - Agreement in need for olecranon osteotomy from fair to moderate ($\kappa_{Traction}$ = 0.28; $\kappa_{2D\text{-}CT}$ = 0.47).
- There was no difference in the inter- or intraobserver reliability of the diagnostic performance parameters measured in experienced fellowship trained surgeons.

Fig. Comparison of (A) injury and (B) traction radiograph of the same fracture demonstrating different characterization capabilities.



Discussion

- Traction radiographs can help the surgeon understand distal humerus fractures and improve reliability of classification and for experienced surgeons they provide similar diagnostic characteristics to 2D-CT
- CT scan may improve reliability of detecting coronal fractures and separate articular fragments especially in less experienced surgeons
- There is fair interobserver reliability in decision for an olecranon osteotomy and poor reliability in decision of treatment reflecting that surgeons rely more on their training biases than on any further information gleaned from traction or CT scanning.