

## Scoring Cortical Healing of Humeral Shaft Fractures Improves Interobserver Reliability

*Anthony Christiano, BA; Abraham Goch, BS; Christopher Burke, MD; Philipp Leucht, MD, PhD; Kenneth Egol, MD; Sanjit Konda, MD; New York University Hospital for Joint Diseases, New York, New York, USA*

**Purpose:** Determining when a fracture of a long bone is united can be difficult. Scoring systems such as the radiographic union score for tibial fractures (RUST) and the radiographic union score for hip (RUSH) have improved interobserver reliability in determining the degree of healing of long bone fractures. The purpose of this study is to determine if a modified RUST score applied to nonoperatively treated humeral shaft fractures can increase interobserver reliability in determining degree of fracture healing.

**Methods:** Three fellowship-trained orthopaedic traumatologists and one fellowship-trained musculoskeletal radiologist scored 50 cases (100 radiographs) of humeral shaft fractures in various stages of healing using a modified RUST scoring system called the Radiologic Humerus Union Measurement (RHUM). All observers were blinded to time from injury. After a 4-week washout period, observers again scored the same cases. Cases were presented in random order with each attempt. Observers classified each fracture as either healed or not healed based on the combination of the two radiographs for the case. Inter- and intraobserver reliability of the RHUM scoring system applied to humeral shaft fractures were determined using an intraclass correlation coefficient (ICC). Interobserver reliability of determining if a fracture is healed was calculated using Cohen's kappa ( $\kappa$ ) statistics.

**Results:** ICC showed almost perfect interobserver reliability (ICC 0.838, ICC 95% confidence interval [CI] 0.765 to 0.896) and intraobserver reliability (ICC range 0.822 to 0.948) with applying the RHUM scoring system to humeral shaft fractures. Cohen's kappa showed substantial agreement between observers in determining fracture healing ( $\kappa = 0.647$ ).

**Conclusion:** The RHUM score applied to humeral shaft fractures showed greater interobserver reliability than overall perception of healing. This is the first time that a cortical scoring system has been shown to have excellent interobserver reliability in a long bone fracture that was not treated operatively. The RUST score applied to the humerus may allow for orthopaedic surgeons to predict healing of humeral shaft fractures, as has been shown for the RUST score in the tibia.