An Image Processing Method for Assessing Bone Healing: A Pilot Study

Francisco Gomez Alvarado, BA*; Mikias Neguisse, BS; Madeline Mackechnie, MA; Emil H. Schemitsch, MD, FIOTA; Theodore Miclau, III, MD, FIOTA*

Purpose: The evaluation of fracture healing is crucial for clinical decision-making and determination of clinical research outcomes. Variability still exists among surgeons determining the state of bone union, and the mRUST (modified Radiographic Union Score for Tibial fractures) score was developed as a validated measure. mRUST scores, which align closely with relevant clinical outcomes, assign values to the cortices of orthogonal anteroposterior (AP) and lateral radiographs based on fracture line, increased bone callus, and bridging. However, limitations due to differences in a rater's expertise still exist. In this pilot study, we seek to compare human-assigned mRUST scores to a measure created through an image processing method.

Methods: 37 AP and lateral radiographs of fractured tibia were obtained from a prospective observational study of adults with open tibia fractures. A human-assigned mRUST score and associated time point of 6, 12, 26, and 52 weeks post-injury were linked to each cortex per radiograph. The Scikit-Image Machine learning package was used in Python 3.10 to process images. For simplicity, we referred to the output of the algorithm as the "opacity index". Statistical analysis was performed using the statsmodels library in Python 3.6.9.

Results: From 37 original radiographs, the corresponding 148 anterior, posterior, medial, and lateral cortices with associated time points and mRUST scores were successfully extracted for analysis. Multinomial logistic regression revealed statistically significant differences in opacity index between mRUST scores 1 and 3 ($P = 0.043$) and 1 and 4 ($P<0.001$). No statistically significant difference was detected between mRUST scores of 1 and 2 ($P = 0.235$). The multinomial logistic regression model converged and adequately fit opacity index to mRUST scores (log likelihood ratio [LLR] $P<0.00001$).

Conclusion: This pilot study reports the use of image processing to create an objective, computer-generated value that suggests correlation with mRUST scores. The study highlights the potential to create an objective tool to measure fracture healing, independent of user expertise. A larger sample size, in addition to image processing algorithmic optimization, is necessary to validate this model.