Can We Predict Distal Radius Fracture Instability? Recognition of Fracture Stability in Distal Radial Fractures on Radiographs: Diagnostic Accuracy of an Artifical Intelligence Algorithm (Convolutional Neural Network (CNN)) to Predict Loss of Threshold Alignment

Britt Barvelink, MD; *Koen Daan Oude Nijhuis, BS; Sanne Floor Hoeksema, BS; Charlotte Louise Laane, MD; Frank Ijpma, MD, PhD; Zhibin Liao, PhD; Ruurd Jaarsma, FRACS, MD; Mathieu Wijffels, MD, PhD; Joost Colaris; Job N. Doornberg, MD, PhD; Machine Learning Consortium*
Erasmus Medical Center, Rotterdam, NETHERLANDS

Purpose: It is challenging to accurately predict the risk of secondary displacement of distal radius fractures (DRFs) after successful closed reduction. One could argue that fracture stability is currently "guesstimated", and is limited by surgeons' bias. In general, surgical reduction and fixation is advised when DRFs are considered "unstable" (Figure 1). Artificial Intelligence (AI) has proven helpful in fracture recognition and classification. Subsequently, we aim to deploy AI to answer the clinically relevant question: Can we develop a Convolutional Neural Network (CNN) that predicts fracture instability on radiographs of DRFs?

Methods: Radiographs of 492 patients with conservatively treated DRFs were retrospectively collected in 2 trauma centers. All radiographs (trauma, postreduction, follow-up) were assessed for acceptable alignment defined by American Academy of Orthopaedic Surgeons (AAOS) International Guidelines. Fractures were classified as unstable, when secondary displacement occurred on follow-up radiographs according to AAOS guidelines, serving as ground truth for the label "unstable". This resulted in 2 categories: 349 stable and 145 unstable fractures. Training and test sets were created and 3 CNN algorithms were developed, based on (1) trauma radiographs, (2) postreduction radiographs, and (3) both trauma and postreduction radiographs. Diagnostic accuracy was calculated according to standard formulas.

Results: Accuracy of the first CNN trained on trauma radiographs was 77% (95% confidence interval [CI] 3.94%); of the second, 73% (95% CI 2.89%); and third, 75% (95% CI 4.00%).

Conclusion: Our International Machine Learning Consortium developed a CNN to predict distal radius fracture instability with an accuracy of 77%. This CNN has great potential to empower both surgeons and patients with personalized risk stratification for fracture instability in the treatment of DRFs. Based on these promising results, we are currently improving accuracy with larger data sets, and will prospectively validate diagnostic performances.