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

Koen Daan Oude Nijhuis[1], Britt Barvelink, Jasper Prijs, Charlotte Louise Laane, RJ Van Luit, Sanne Floor Hoeksema, Zhibin Liao, Yang Zhao, Frank Ijpma, Ruurd Jaarsma[2], Mathieu Wijffels, Joost Colaris, Job N Doornberg[1]
[1]University Medical Centre Groningen, [2]Flinders University and Flinders Medical Centre

INTRODUCTION:

Distal radius fractures (DRFs) are very common in adults. It is challenging to accurately predict the risk of secondary displacement of 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). Unfortunately, up to 64% of reduced DRFs appear to be unstable, which results in fracture redisplacement during the first treatment weeks. Defining the right treatment strategy in an early phase, especially in the emergency room, is difficult but essential.

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 two Trauma Centers. All radiographs (trauma, post-reduction, follow up) were assessed for acceptable alignment defined by AAOS International Guidelines.

Fractures were classified as unstable, when secondary displacement occurred on follow-up radiographs according to AAOS guidelines, serving as the ground truth for the label unstable. This resulted in two categories: 349 stable and 145 unstable fractures.

Training- and test-sets were created and three CNN algorithms were developed, based on 1) trauma radiographs; 2) post-reduction radiographs; 3) both trauma and post-reduction radiographs. Diagnostic accuracy was calculated according standard formulas using a 10-fold Monte Carlo cross-validation.

RESULTS:

Accuracy of the first CNN trained on only trauma radiographs is 71% (95% CI 6.54%). The algorithm trained on post-reduction radiographs is 69% (95% CI 5.62%). The accuracy of the algorithm trained on both trauma and post-reduction radiographs was the highest at 73% (95% CI 4.01%). Figure 2 shows the output of the algorithm on a patient from the test-set.

DISCUSSION AND CONCLUSION:

Our International Machine Learning Consortium developed a CNN to predict distal radius fracture instability with an accuracy of 73%. 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 datasets, and will prospectively validate diagnostic performances.

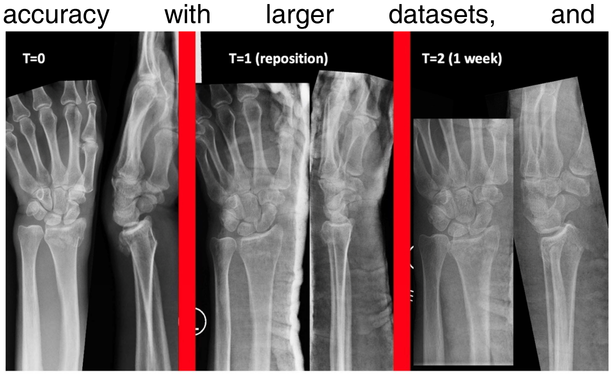

Figure 1: A Distal Radial Fracture that shows secondary dislocation.

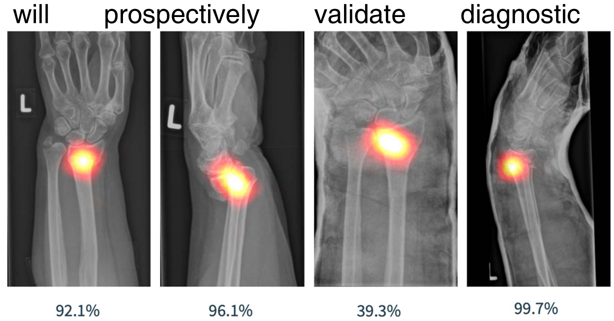

Figure 2: Radiographs from a patient in the test-set. A heatmap is created by the algorithm, visualizing the area the algorithm focuses on. The certainty of the prediction is listed. This fracture is unstable.