Evaluating the Accuracy of Radiograph-Based Machine Learning in Diagnosing Distal Radius Fractures: A Systematic Review and Meta-Analysis

A Msherghi, M Abuajamieh, Z Alawi, M Azlitni, A Fahaid, N Ouakouak, A Alsayed, M Elhadi, CF Beaulieu, M Wintermark

Background

Distal radius fractures (DRFs) are the most common adult bone fractures, accounting for 17.5% of cases
Recent studies suggest that integrating machine learning (ML) into medical imaging can improve the
detection of bone anomalies. [2] . We propose that ML algorithms could significantly enhance DRF image
analysis, especially in emergency departments where quick and accurate diagnosis is crucial. We aim to
evaluate the diagnostic accuracy of radiograph-based ML algorithms in detecting DRFs compared to
expert readings.

Methods

A comprehensive search on PubMed, EMBASE, Scopus, Web of Science, the Cochrane Library,
and IEEE from inception to August 2024 focused on ML's performance in detecting DRFs. The
Quality Assessment of Diagnostic Accuracy Studies (QUADAS-2) tool was used to assess study
quality. A funnel plot indicated symmetry across studies, suggesting no publication bias. Statistical
analyses were done using R (version 4.0.3) with the mada package. This study is registered in
PROSPERO (CRD42024520163).

Results

We screened 8,854 records, of which 10 out of 22 eligible studies were included, encompassing a
total of 2,765 radiographs. Among these, 1,375 radiographs confirmed fractures, and 1,390 showed
non-fractures. The ML algorithms achieved a sensitivity of 92.7% (95% CI: 88.3-95.5%, I^2 = 81.4%)
and a specificity of 93.9% (95% CI: 88.3-96.9%, I^2 = 83.9%). The diagnostic odds ratio was
calculated to be 196.48 (95% CI: 81.15-475.70). The positive likelihood ratio (LR+) was 15.28 (95%
CI: 7.77-30.03), and the negative likelihood ratio (LR-) was 0.08 (95% CI: 0.05-0.12). Nine studies
(90%) clustered in the high-accuracy area of the SROC curve, reflecting a high diagnostic
performance.

Conclusion

Conventional radiograph-based ML algorithms perform comparably to imaging experts in detecting
DRFs, making them a valuable supplementary diagnostic tool. However, due to the observed
heterogeneity across studies, further high-quality, longitudinal trials are needed to develop standard
guidelines for the application of such ML in clinical settings.

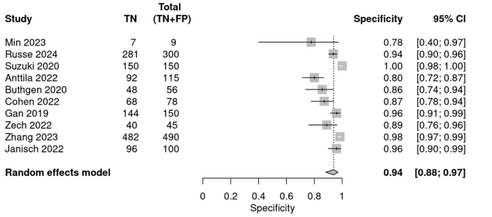
Contact

AAmsherghi@mdanderson.org

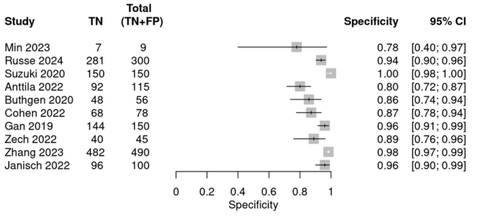
Conflict of interest

No conflict of interest to declare.

Sensitivity



Specificity



SROC



RESEARCH POSTER PRESENTATION DESIGN © 2015
www.PosterPresentations.com