A Retrospective Study on the Performance of an AI Tool in Detecting Vertebral Fragility Fractures Across Multiple CT Protocols

DC Schramm[1], J Cope[2], E Preston[2], R Eckert[2], MK Javaid[3], S Ather[2]

1. Medical Sciences Division, University of Oxford, Oxford, UK
2. Department of Radiology, Nuffield Orthopaedic Centre, Oxford University Hospitals
3. Nuffield Department of Orthopaedics, Rheumatology and Musculoskeletal Sciences, University of Oxford

Introduction

Vertebral fragility fractures (VFFs) are a major concern in aging populations, particularly as they are the most common osteoporotic fractures. These fractures are key predictors of future osteoporotic fractures and are associated with increased morbidity and mortality.[1-2] HealthVCF is an AI tool, is designed to assist radiologists in identifying Genant Grade ≥2 VFFs on CT scans.[3]

Figure 1: Genant Classification (adapted from Wang et al.)[4]



Methods

This retrospective study reviewed CT scans from patients aged 70 or older at Oxford University Hospitals between 17/10/23 and 07/11/23. Each CT scan was manually reviewed for VFFs. Only VFFs grade ≥2 based on the Genant classification were considered radiologist 'positive' fractures.

Inclusion Criteria

- Patients aged 70 or older
- CT scans from the chest, abdomen, and pelvis (CTCAP), chest (CTC), abdomen and pelvis (CTAP), pulmonary angiogram (CTPA), and kidneys, ureters, and bladder (CTKUB)

Exclusion Criteria

- Trauma-related CT scans
- Known bony metastases, myeloma, or significant spinal deformities

References

1.Schousboe JT. Epidemiology of Vertebral Fractures. *Journal of Clinical Densitometry*. 2016;19(1):8-22. doi:10.1016/j.jocd.2015.08.004
2. Melton III LJ, Atkinson EJ, Cooper C, O'Fallon WM, Riggs BL. Vertebral Fractures Predict Subsequent Fractures. *Osteoporos Int*. 1999;10(3):214-221. doi:10.1007/s001980050218
3. Bendtsen MG, Hitz MF. Opportunistic Identification of Vertebral Compression Fractures on CT Scans of the Chest and Abdomen, Using an AI Algorithm, in a Real-Life Setting. *Calcif Tissue Int*. 2024;114(5):468-479. doi:10.1007/s00223-024-01196-2
4. Wáng YX, Ruiz Santiago F, Deng M, Nogueira-Barbosa M. Identifying osteoporotic vertebral endplate and cortex fractures. Quantitative Imaging in Medicine and Surgery. 2017;7:555-591. doi:10.21037/qims.2017.10.05

Objective

To assess the diagnostic performance of HealthVCF in detecting vertebral fragility fractures across multiple CT protocols (CTCAP, CTC, CTAP, CTPA, and CTKUB).

Why?

The use of AI in radiology is expanding, and it is important to validate the performance of AI tools in real-world settings across diverse imaging protocols to ensure consistent and accurate detection of critical conditions like VFFs.

HealthVCF had an overall sensitivity of 70.3% and specificity of 96.8% in detecting vertebral fragility fractures across multiple CT protocols.

Results

537 CT scans were analysed. These included 103 CTCAPs, 99 CTCs, 127 CTAPs, 103 CTKUBs, and 105 CTPAs. 18.8% of patients were diagnosed with Genant grade ≥2 VFFs.

Confusion Matrix

Predicted \ Actual	Radiologist Positive	Radiologist Negative
HealthVCF Positive	71	14
HealthVCF Negative	30	422

DETAILS

Sensitivity	Specificity	PPV	NPV	Accuracy
70.3%	96.8%	83.5%	93.4%	91.8%
95% CI: 60.8-78.3	95% CI: 94.7-98.1	95% CI: 74.2-89.9	95% CI: 90.7-95.3	95% CI: 89.2-93.8

HealthVCF demonstrated a high specificity and moderate sensitivity for VFFs. Notably, HealthVCF missed 30 fractures, including 15 each of grade 2 and grade 3 VFFs, while generating 14 false positives, most of which were associated with degenerative disease.

Sensitivity and Specificity by CT Protocol

Sensitivity and specificity varied by protocol, with the highest sensitivity in CT chest (80.8%) and the highest specificity in CTPA (97.7%). CTCAP and CTKUB showed lower sensitivity, possibly due to differing protocol setting and motion artifact.

