Artificial Intelligence Enabled Vertebral Fracture Identification for Fracture Liaison Services: opportunities and challenges- The ADOPT Study

Chappell DDG.[1]; Sanders V[1]; Vindlacheruvu. M[2]. Boylan JM[2]; Turton J[3]; Sampson M[4]; Guy A[4]; Eckert R[5]; Curtis E[6]; Stone M[2]; Cooper F[7]; Kimmel Y[8]; Poole KES[1]; Sahota O[9]; Javaid MK[10]

[1]Department of Rheumatology, Cambridge University Hospitals NHS Foundation Trust, Cambridge, UK

[2]Department of Trauma and Orthopaedics, Cambridge University Hospitals NHS Foundation Trust, UK

[3]Cardiff and Vale University Health Board, Cardiff, UK

[4]University Hospitals Southampton Foundation Trust, Southampton, UK

[5]Oxfordshire Fracture Prevention Service, Oxford University Hospitals NHS Foundation Trust, Oxford, UK

[6]MRC Lifecourse Epidemiology Centre, University of Southampton, Southampton, UK

[7]Royal Osteoporosis Society, UK

[8]NanoxAI, Nano-X Imaging Ltd

[9]Nottingham University Hospitals NHS Trust, Nottingham, UK

[10]NDORMS, University of Oxford, Oxford, UK

Objective: To assess the effectiveness of artificial intelligence (AI) in enhancing the detection and management of vertebral fragility fractures (VFFs), which are often under diagnosed, leading to missed opportunities for secondary fracture prevention. We evaluated the integration of AI into Fracture Liaison Services (FLSs) across UK public hospitals.
Methodology: AI was implemented in four NHS hospitals in England and Wales using HealthVCF algorithm, which was developed to detect moderate to severe thoraco-lumbar VFFs from routine CT scans. After local information governance approval, a shadow test comprising 500 consecutive CT scans from 2017 was conducted at each site for clinical validation before routine

clinical use. FLS-A opted for synchronous augmented live reporting by radiologists who used a standard shortcode to identify reports to be exported to the FLS. The remaining three FLSs forwarded eligible scans for asynchronous local clinical confirmation and FLS management. FLSs used local improvement teams to review and improve their performance.

Outcomes: Post-shadow testing, FLS-A chose the high specificity (98.0%) mode to minimize false positive scans, while the other sites chose a balanced mode with a higher sensitivity (87% in the balanced vs 59% in the high sensitivity mode). The table illustrates the number of AI-reviewed scans, those flagged with potential VFFs, locally reviewed and clinically confirmed over approximately three months. Improvement teams identified common and distinct issues including: criteria to filter out patients not requiring FLS assessment, improving invitation letters to patients, monitoring and digital tools to reduce administration time and improve patient safety.

Conclusion: The incorporation of AI in identifying VFFs significantly increased patient identification rates. It also underscored specific areas requiring refinement if FLSs are to fully realize anticipated improvements from AI on patient outcomes.

Table: Number of CT scans analysed, flagged, clinically reviewed and confirmed over approximately 3 months relative to 2022 FLSDB submitted reports

Sites	FLS-A	FLS-B	FLS-C	FLS-D
AI CT analyzed scans	3298	20,239	3654	9,429
AI CT flagged scans	410 (12.4%)	5,876 (29.0%)	1,484 (40.6%)	2,936 (31.1%)
Clinically reviewed scans	410	3,575	1,809	1,655
Clinically positive scans	221	1,367	282	646
Increase from 2022 FLSDB submitted spine records[*] (%)	42%	2,539%	354%	447%
Percentage of 2022 FLSDB total records submitted[*]	33.7%	65.2%	36.7%	31.6%

Legend:

FLS-A implemented augumented synchronous reporting while the other FLSs implemented forwarding of eligible scans for asynchronous clinical confirmation

[*]https://www.fffap.org.uk/FLS/charts.nsf/benchmarks?ReadForm&yr=2022&vw=BYO&org1=SGH,BRD,ADD,LLD