Evaluating the impact of artificial intelligence-assisted image analysis on the diagnostic accuracy of front-line clinicians in detecting fractures on plain X-rays (FRACT-AI): A multicase multireader study

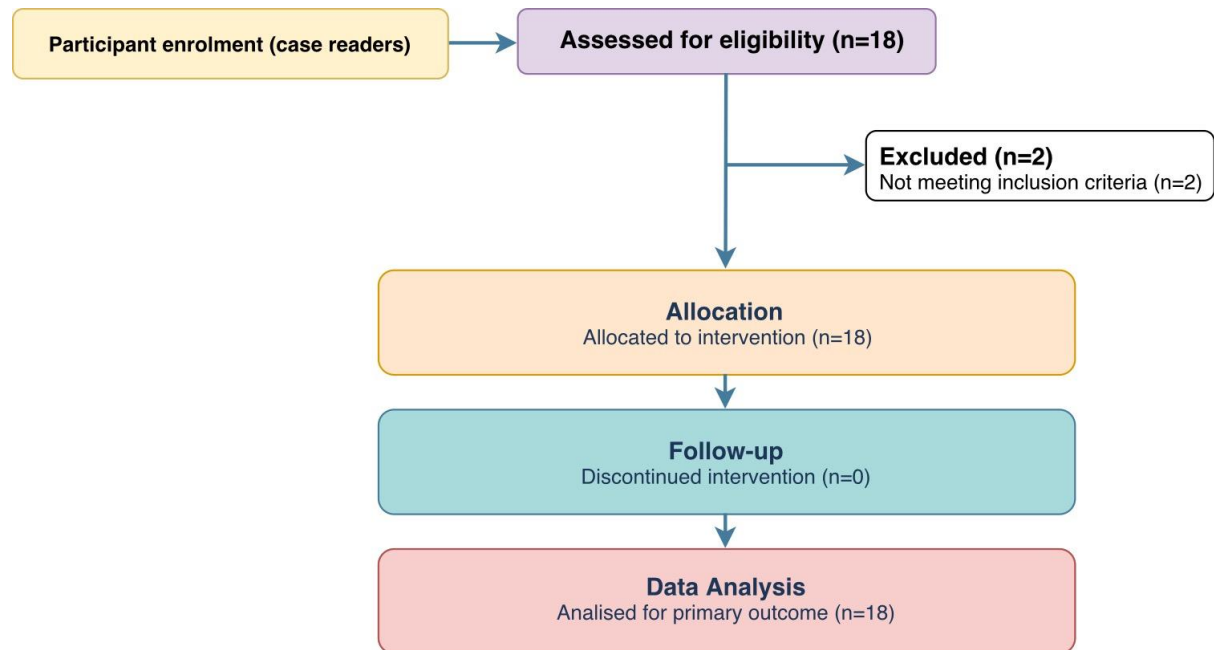**Figure 1. Participants flowchart**

Table 1. Baseline characteristics

Specialty	Senior (>10 yrs)	Middle grade (5–10 yrs)	Junior (<5 yrs)
Emergency Physicians	1	1	1
Trauma & Orthopaedic Surgeons	1	1	1
Emergency Nurse Practitioners	1	1	1
Physiotherapists	1	1	1
General Radiologists	1	1	1
Reporting Radiographers	1	1	1
Total participants	6	6	6

Outcome Measures: Tables 2 to 8

Finding	Sensitivity % (95% CI)	Specificity % (95% CI)
Abnormality	87.5 (82.5-91.5)	87.7 (83.1-91.4)
- Fracture	85.1 (79.8-89.4)	89.7 (85.5-93)
- Dislocation	66.7 (22.3-95.7)	96.6 (94.5-98)

Table 2. Analysis of the algorithm performance against ground truth on a per case basis

Finding	Sensitivity % (95% CI)	Specificity % (95% CI)
Abnormality	79.8 (68.6, 77.0)	91.7 (89.4, 93.6)
- Fracture	70.8 (66.3, 74.9)	93.0 (90.8, 94.7)
- Dislocation	50.0 (23.7, 76.3)	98.3 (97.4, 99.0)

Table 3. Analysis of the algorithm performance against ground truth on a per image basis

Analysis (Pooled reader, per case)	Without AI % (95% CI)	With AI % (95% CI)	Difference % (95% CI)	p value
AUC (Abnormality)	0.883 (0.858, 0.907)	0.921 (0.901, 0.942)	+0.038 (0.021, 0.057)	<0.001
Sensitivity (Abnormality)	82.8% (78.8, 86.8)	86.7% (82.7, 90.6)	+3.9 (1.7, 6)	<0.001
Specificity (Abnormality)	82.9% (78.4, 87.5)	90.4% (87.7, 93.1)	+7.5 (3.6, 11.3)	<0.001
AUC (Fracture)	0.878 (0.852,0. 903)	0.917 (0.896, 0.938)	+0.039 (0.021, 0.057)	<0.001
Sensitivity (Fracture)	81.5 (77.3, 85.7)	85.5 (81.5, 89.5)	+4.0 (1.8, 6.3)	<0.001
Specificity (Fracture)	84.2 (79.8, 88.5)	90.9 (88.3, 93.6)	+6.7 (3.3, 10.3)	<0.001
AUC (Dislocation)	0.971 (0.952, 0.99)	0.936 (0.865, 1.00)	-0.035 (-0.095, 0.026)	0.257
Sensitivity (Dislocation)	86.1 (76.1, 96.1)	82.4 (64.7, 100)	-3.7 (-21.4, 14)	0.682
Specificity (Dislocation)	98.1 (97.4, 98.9)	98.7 (97.9, 99.5)	+0.6 (-0.05, 1.2)	0.072

Table 4. Pooled analysis of per case overall reader performance

Analysis (Pooled reader, per image)	Without AI % (95% CI)	With AI % (95% CI)	Difference % (95% CI)	p value
Sensitivity (Abnormality)	61.8 (54.6, 69.0)	60.0 (49.1, 70.8)	-1.8 (-8.2, 4.5)	0.55
Specificity (Abnormality)	90.6 (88.8, 92.3)	94.4 (92.8, 96.0)	+3.8 (2.0, 5.7)	<0.001
Sensitivity (Fracture)	61.1(54.0, 68.3)	59.4 (48.6, 70.2)	-1.7 (-8.02, 4.63)	0.58
Specificity (Fracture)	91.4 (89.7, 93.1)	94.8 (93.3, 96.3)	+3.4 (1.60, 5.11)	<0.001
Sensitivity (Dislocation)	56.7 (40.9, 72.5)	50.0 (32.3, 67.7)	-6.7 (-17.4, 4.1)	0.22
Specificity (Dislocation)	98.8 (98.3, 99.3)	99.2 (98.8, 99.7)	+0.4 (0.09, 0.75)	0.012

Table 5. Pooled analysis of per image overall reader performance

Analysis (per case)	Without AI % (95% CI)	With AI % (95% CI)	Difference (95% CI)	p value
AUC (Abnormality) EM	83.6 (78.4, 88.9)	91.4 (85.5, 97.2)	+7.8 (3.0, 2.5)	0.014
Sensitivity (Abnormality) EM	75.1 (59.3, 91.0)	81.5 (61.0, 100)	+6.3 (2.7, 10.0)	0.007
Specificity (Abnormality) EM	81.2 (67.8, 94.6)	91.9 (83.6, 100)	+10.7 (-3.9, 25.3)	0.092
AUC (Abnormality) T&O	86.1(79.6, 92.6)	91.2 (87.5, 95.0)	+5.1 (2.6, 7.8)	0.001
Sensitivity (Abnormality) T&O	82.3 (78.6, 86.0)	86.2 (82.3, 90.1)	+3.9 (-1.0, 8.8)	0.104
Specificity (Abnormality) T&O	74.4 (44.6, 100)	88.2 (83.4, 92.9)	+13.8 (-10.5, 38.1)	0.143
AUC (Abnormality) ENP	86.3 (83.3, 89.4)	91.9 (89.6, 94.2)	+5.6 (2.9, 8.3)	0.002
Sensitivity (Abnormality) ENP	83.3 (74.1, 92.6)	87.8 (82.9, 92.7)	+4.5 (-3.4, 12.3)	0.188
Specificity (Abnormality) ENP	82.8 (61.7, 100)	88.3 (82.0, 94.6)	+5.5 (-12.4, 23.3)	0.34
AUC (Abnormality) Radiology	92.1 (90.1, 94.1)	93.8 (91.8, 95.9)	+1.7 (-2, 5.5)	0.183
Sensitivity (Abnormality) Radiology	85.3 (81.8, 88.9)	88.9 (85.5, 92.4)	+3.6 (0.6, 6.6)	0.032
Specificity (Abnormality) Radiology	90.4 (82.7, 98.1)	93.3 (90.9, 95.7)	+2.9 (-6.1, 11.8)	0.31
AUC (Abnormality) Physiotherapy	89.5 (85.9, 93.1)	91.9 (89.0, 94.7)	+2.4 (0.43, 4.3)	0.017
Sensitivity (Abnormality) Physiotherapy	83.3 (74.1, 92.6)	87.8 (82.9, 92.7)	+4.5 (-3.36, 12.27)	0.188
Specificity (Abnormality) Physiotherapy	82.8 (61.7, 100)	88.3 (82.0, 94.6)	+5.5 (-12.4, 23.3)	0.34
AUC (Abnormality) Radiography	91.9 (89.9, 93.9)	92.5 (90.3, 94.6)	+0.6 (0.047, 1.08)	0.033
Sensitivity (Abnormality) Radiography	88.2 (83.7, 92.8)	89.2 (84.7, 93.8)	+1.0 (-1.0, 3.0)	0.261
Specificity (Abnormality) Radiography	89.7 (83.8, 95.5)	91.2 (86.7, 95.6)	+1.5 (-5.4, 8.4)	0.47

Table 6. Per case subgroup analysis for abnormality by specialty

Analysis	Without AI % (95% CI)	With AI % (95% CI)	Difference (95% CI)	p value
AUC (Abnormality) Senior	86.4 (79.1, 93.8)	91.1 (87.5, 94.6)	+4.63 (0.76, 8.5)	0.03
Sensitivity (Abnormality) Senior	78.2 (66.3, 90.1)	81.9 (70.8, 93.0)	+3.66 (1.66, 5.67)	0.002
Specificity (Abnormality) Senior	86.7 (80.0, 93.3)	92.7 (88.1, 97.4)	+6.06 (-0.153, 12.28)	0.054
AUC (Abnormality) Middle	87.3 (81.9, 92.7)	92.0 (89.3, 94.6)	+4.63 (0.557, 8.71)	0.034
Sensitivity (Abnormality) Middle	84.6 (76.7, 92.5)	87.5 (82.2, 92.8)	+2.91 (-1.1, 6.92)	0.112
Specificity (Abnormality) Middle	77.5 (56.1, 98.9)	89.1 (85.3, 92.9)	+11.6 (-6.66, 29.79)	0.139
AUC (Abnormality) Junior	87.3 (81.7, 92.9)	93.2 (90.7, 95.8)	+5.94 (1.06, 10.81)	0.028
Sensitivity (Abnormality) Junior	81.2 (77, 85.5)	87.8 (84.0, 91.7)	+6.57 (4.26, 8.88)	<0.001
Specificity (Abnormality) Junior	79.7 (72.2, 87.1)	90.5 (86.7, 94.3)	+10.8 (4.63, 17.01)	0.008

Table 7. Per case subgroup analyses for abnormality by seniority/grade

Reader Group	Mean interpretation time without AI (s)	Mean interpretation time with AI (s)	p value
Pooled	43.0	42.5	0.835

Table 8. Pooled analysis of mean interpretation time per case

Adverse Events

There were no adverse events associated with this study.