

## Appendix 1

Confusion matrix demonstrates the average true positive, true negative, false positive, and false negative values for the artificial intelligence (AI) model, the expert group, and the non-expert group on the independent test data set. The average positive predicted value (PPV) and negative

predicted value (NPV) of the model were 78.1% (24.6/31.5) and 97.8% (18.1/18.5), respectively on the anteroposterior (AP) view. While on the paired views, the average PPV (11.4/11.5) and NPV (8.9/9.5) of the model were 99.1% and 93.7%, respectively. The average NPV for the non-expert group was 76.1% (21.6/28.4) and 75.3% on the AP and paired views, respectively.

**Table S1** Confusion matrix for the AI model, the expert group and the non-expert group on the AP view

Predicted	Actual	
	Normal	Fracture
AI		
Normal	18.1	0.4
Fracture	6.9	24.6
Experts		
Normal	21.5	1.1
Fracture	3.5	23.9
Non-experts		
Normal	21.6	6.8
Fracture	3.4	18.2

AP, anteroposterior; AI, artificial intelligence.

**Table S2** Confusion matrix for the AI model, the expert group, and the non-expert group on the paired AP and lateral view

Predicted	Actual	
	Normal	Fracture
AI		
Normal	8.9	0.6
Fracture	0.1	11.4
Experts		
Normal	7.9	0.9
Fracture	1.1	11.1
Non-experts		
Normal	7.3	2.4
Fracture	1.7	9.6

AP, anteroposterior; AI, artificial intelligence.