

Table S1 MAE & 95% CI of AI and reviewers in RUS

Group	Sample size	MAE of reviewer-gold	95% CI of viewer-gold	MAE of AI-gold	95% CI of AI-gold
Age (years)					
1–6	172	0.01	(–0.03, 0.05)	–0.02	(–0.06, 0.03)
7–12	367	–0.01	(–0.04, 0.026)	–0.06	(–0.09, –0.03)
13–18	178	0.02	(–0.03, 0.07)	–0.15	(–0.20, –0.11)
19–20	27	–0.04	(–0.11, 0.02)	–0.14	(–0.20, –0.08)
Gender					
Male	378	0.03	(–0.00, 0.06)	–0.05	(–0.08, –0.02)
Female	366	–0.03	(–0.06, 0.01)	–0.10	(–0.13, –0.07)

MAE, mean absolute error; CI, confidence interval; AI, artificial intelligence; RUS, radius, ulna and short bones.

Table S2 MAE & 95% CI of AI and reviewers in Carpal

Group	Sample size	MAE of reviewer-gold	95% CI of viewer-gold	MAE of AI-gold	95% CI of AI-gold
Age (years)					
1–6	172	0.05	(0.01, 0.08)	–0.17	(–0.20, –0.13)
7–12	367	0.03	(–0.02, 0.07)	–0.07	(–0.12, –0.02)
13–18	178	–0.08	(–0.17, 0.01)	–0.35	(–0.44, –0.26)
19–20	27	–0.27	(–0.49, –0.06)	–0.94	(–1.22, –0.66)
Gender					
Male	378	0.03	(–0.02, 0.08)	–0.12	(–0.17, –0.07)
Female	366	–0.03	(–0.08, 0.01)	–0.24	(–0.29, –0.19)

MAE, mean absolute error; CI, confidence interval; AI, artificial intelligence.

Table S3 The MAE & 95% CI of BAA in different manufacturers in RUS

Manufacturer	Sample size	MAE of reviewer-gold	95% CI of viewer-gold	MAE of AI-gold	95% CI of AI-gold
Canon Inc.	43	0.07	(–0.02, 0.16)	–0.02	(–0.10, 0.06)
Carestream Health	107	–0.05	(–0.11, 0.01)	–0.16	(–0.23, –0.10)
GE HealthCare	63	0.06	(–0.03, 0.15)	–0.04	(–0.12, 0.05)
Kodiak	173	–0.03	(–0.08, 0.02)	–0.08	(–0.13, –0.03)
Philips Medical	195	0.02	(–0.02, 0.07)	–0.02	(–0.06, 0.02)
Siemens	96	0.02	(–0.04, 0.09)	–0.10	(–0.15, –0.04)

MAE, mean absolute error; CI, confidence intervals; BAA, bone age assessment; RUS, radius, ulna and short bones; AI, artificial intelligence.

Table S4 The MAE & 95% CI of BAA in different manufacturers in Carpal

Manufacturer	Sample size	MAE of reviewer-gold	95% CI of viewer-gold	MAE of AI-gold	95% CI of AI-gold
Canon Inc.	43	-0.09	(-0.23, 0.05)	-0.05	(-0.21, 0.11)
Carestream Health	107	-0.08	(-0.17, 0.01)	-0.13	(-0.22, -0.04)
GE HealthCare	63	0.06	(-0.04, 0.15)	0.09	(-0.01, 0.19)
Kodiak	173	-0.02	(-0.09, 0.06)	-0.03	(-0.11, 0.05)
Philips Medical	195	0.02	(-0.04, 0.09)	-0.44	(-0.51, -0.37)
Siemens	96	0.03	(-0.06, 0.13)	-0.28	(-0.37, -0.19)

MAE, mean absolute error; CI, confidence interval; BAA, bone age assessment; AI, artificial intelligence.

Table S5 Kappa test between gold standard and AI or reviewers

Bones	AI	Reviewer 1	Reviewer 2	Reviewer 3	Reviewer's mean
TW3-Carpal					
Capitate	0.56	0.66	0.68	0.69	0.68
Hamate	0.68	0.76	0.72	0.74	0.74
Triquetrum	0.69	0.79	0.80	0.72	0.77
Lunate	0.63	0.74	0.77	0.75	0.75
Scaphoid	0.70	0.78	0.82	0.77	0.79
Trapezium	0.66	0.78	0.80	0.76	0.78
Trapezoid	0.65	0.76	0.76	0.74	0.76
TW3-RUS					
Radius	0.58	0.73	0.76	0.75	0.75
Ulna	0.74	0.83	0.86	0.82	0.84
Metacarpal I	0.69	0.78	0.83	0.82	0.81
Metacarpal III	0.65	0.79	0.81	0.77	0.79
Metacarpal V	0.62	0.78	0.81	0.78	0.79
Proximal phalange I	0.74	0.84	0.85	0.83	0.84
Proximal phalange III	0.73	0.84	0.84	0.85	0.84
Proximal phalange V	0.74	0.84	0.84	0.83	0.84
Middle phalange III	0.81	0.83	0.88	0.86	0.86
Middle phalange V	0.70	0.78	0.79	0.79	0.79
Distal phalange I	0.69	0.78	0.81	0.78	0.79
Distal phalange III	0.67	0.74	0.83	0.79	0.79
Distal phalange V	0.71	0.75	0.80	0.79	0.78

AI, artificial intelligence; TW3, Tanner-Whitehouse 3; RUS, radius, ulna and short bones.