

Chest X-ray scoring of COVID-19: Inter observer variability&Application for artificial intelligence

순천향대학교 서울병원 내과¹, 순천향대학교 서울병원 감염내과², 순천향대학교 부천병원 임상시험센터 의학통계상담실³, 한동대학교 기계제어공학부⁴, 순천향대학교 구미병원 영상의학과⁵, 서울벤처대학원대학교 융합산업학과⁶, 순천향대학교 구미병원 호흡기알레르기내과⁷

김민정¹, 박세윤², 문지은³, 문길래⁴, 임한혁⁵, 양승부⁵, 황찬규⁶, 김명신⁷, 임건일⁷, 황현규⁷

Background/Aims: The advent of SARS-CoV-2 led to a pandemic with confirmed cases of over 15 million. However, information on risk factors which determine severity and prognosis of the disease is scarce, in particular, in medically deprived areas where the use of computed tomography is limited. Therefore we conducted a pilot study to evaluate whether the initial chest x-ray image can predict the patient's disease course, and furthermore whether it could be applied clinically by teaching artificial intelligence.

Methods: Chest x-ray (CXR) images of COVID-19 patients (total 21 patients) were collected who were admitted to Soonchunhyang University Hospital Seoul between February 2020 and May 2020 (13 patients) and patients who had positive COVID-19 test results at Soonchunhyang University Hospital Gumi between February 2020 and 21st July 2020 (8 patients). Two interpreters (Interpreter 1, a radiologist; Interpreter 2, a physician) scored each x-ray images with CXR scoring system introduced by A. Borghesi et al. In the CXR scoring system, a frontal chest radiograph is divided into six zones (from A to F) using two lines. Each zone of CXR is assigned with a score among 0 and 3 depending on severity of infiltration. We analyzed the agreement of each zone between two interpreters using weighted Kappa method and had artificial intelligence learn the CXR scoring system with data of Interpreter 1.

Results: The weighted Kappa result was 0.5714, 0.625, 0.3186, 0.5401, 0.5743, and 0.3576 for each section (Table 1). When divided into two groups of CXR score ≥ 3 and CXR score < 3 , the difference of demographic between the two groups was as follow (Table 2). The difference of symptoms and outcomes of the two groups is as follow (Table 3). We are training our data on artificial intelligence and waiting for sensitivity results.

Conclusions: The CXR scoring may be a useful method to read chest images of COVID-19 and learning this in artificial intelligence will help to quickly determine the severity of COVID-19 patients in the clinical setting. Subsequent large scale study is needed for further validation.

Table 1 Agreement between two interpreters calculated by Weighted Kappa

Table1-1. A zone					Table1-2. B zone					Table1-3. C zone				
CXR score	0	1	2	3	CXR score	0	1	2	3	CXR score	0	1	2	3
0	16	0	0	0	0	16	0	0	0	0	8	1	0	0
1	2	1	0	0	1	3	0	0	0	1	9	1	0	0
2	0	0	0	0	2	0	0	1	0	2	0	1	0	0
3	0	1	2	0	3	0	0	1	0	3	0	0	0	1
Weighted Kappa = 0.5714; 95% confidence interval 0.3525-0.7904					Weighted Kappa = 0.625; 95% confidence interval 0.2919-0.9581					Weighted Kappa = 0.3186; 95% confidence interval 0-0.7516				
Table1-4. D zone					Table1-5. E zone					Table1-6. F zone				
CXR score	0	1	2	3	CXR score	0	1	2	3	CXR score	0	1	2	3
0	18	1	0	0	0	14	0	0	0	0	8	0	0	0
1	1	0	0	0	1	5	0	0	0	1	8	0	0	0
2	0	0	0	0	2	0	0	0	0	2	2	1	1	0
3	0	0	1	0	3	0	0	1	1	3	0	0	0	1
Weighted Kappa = 0.5401; 95% confidence interval 0.1041-0.9762					Weighted Kappa = 0.5743; 95% confidence interval 0.210-0.9387					Weighted Kappa = 0.3576; 95% confidence interval 0.0194-0.6959				

Table 2 Demographics of two groups [CXR score ≥ 3 and CXR score < 3]

	CXR score < 3 (n=13)	CXR score ≥ 3 (n=8)
Age (mean)	32.62	44.25
Sex (female,%)	6 (46.15%)	2 (25%)
Underlying disease		
DM	0	1
Hypertension	1	3
Cancer	0	1
AIDS	2	0
Others	0	2 (schizophrenia, dyslipidemia)

Table 3 Symptoms and outcomes of two groups [CXR score ≥ 3 and CXR score < 3]

	CXR score < 3 (n=13)	CXR score ≥ 3 (n=8)
Symptoms		
asymptomatic	1	0
cough	7	5
sputum	2	2
dyspnea	1	1
fever	5	3
myalgia	5	0
ageusia	0	2
anosmia	0	3
sore throat	3	1
headache	1	2
Others	3(diarrhea,chill,chest pain)	4 (chill, nasal obstruction, rhinorrhea, lt side weakness and dysarthria)
Outcomes		
presence of follow-up chest x-ray	8	6
Worsening of CXR(CXR score increase ≥ 3)	1	1
Percentage (%)	12.50%	16.67%
ICU admission	0	0
Ventilator care	0	1
Death	0	1