Artificial Intelligence Solution for Identification of Actionable Abnormalities on Chest Radiographs: A Retrospective Case-control Study

Moon Young Kim[1], Jin Young Yoo[2], Yunhee Jang[3], Junghee Hong[4], Sanghyup Lee[5], Kwang Nam Jin[1]

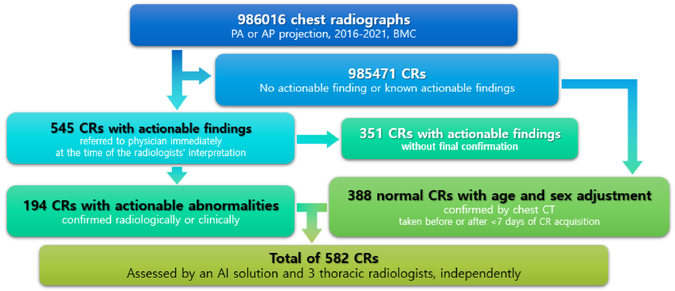
[1]SMG-SNU Boramae Medical Center, Korea, Republic of, Department of Radiology, Seoul National University College of Medicine, Seoul, Republic of Korea
[2]Chungbuk National University Hospital, Korea, Republic of
[3]Sejong General Hospital, Korea, Republic of
[4]Keimyung University Dongsan Medical Center, Korea, Republic of
[5]Lunit, Korea, Republic of

1. BACKGROUND

- The prevalence of actionable findings in chest radiographs(CRs) **32.5%** in a large university hospital
  ER: 4.6% = 46% x 10%, Routine care: **27.9%** = 31% x 90% of total CRs in an institution
- **Chest radiograph(CR)**: Basic imaging for thoracic disease 26% of all diagnostic radiology tests.
  Increasing number in a variety area of medical care
  Practical limitation, interpretation priority, full-time availability of radiologists
- Need for a **triage**(to differentiate CRs with **"actionable" findings** from CRs with nonactionable ones before radiologists start to read CRs)
- **Can AI be a triage in a whole general clinical care?**
  → improve the quality of radiologic interpretation and overall turnaround time

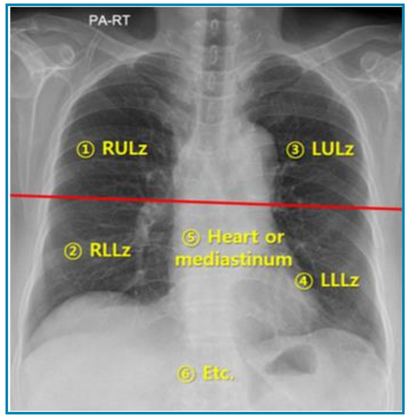
2. MATERIALS AND METHOD

- **Deep Learning Algorithm**(Lunit INSIGHT CXR, version 2.0.2)
  Detection of 10 findings in CRs
  Nodule, Consolidation, Pneumothorax, Pleural effusion, Cardiomegaly, Fibrosis, Pneumoperitoneum, Mediastinal widening, Calcification, Atelectasis
  - Positive result: abnormality score of **15% or more**



- **Figure 1. A flow chart for selecting CXRs for the dataset.**

- **Statistical Analysis**
  Performances of AI solution and reviewers
  - Area under the receiver operating characteristic curve(AUC)
  - Sensitivity, specificity
  - Detection yield: Proportion of true-positive identification among the entire examination
  - false-referral rate(FRR): Proportion of false-positive identification among all examinations
  SPSS 26.0, Rex 3.6.0, p <0.05 for statistical significance



- **Figure 2. Image interpretation by Radiologists**: Independent assessment by three thoracic radiologists(7-10 years of experience) to localize actionable findings

3. RESULTS

- **Table 1. Characteristics of the study population and the dataset**

	Male (63.9%)	Age (years)	CRs with AP projection	CRs from inpatients	CRs from ER	AI score
Total Subjects (n = 582)	372	63.0±17.7	158 (27.1%)	70 (12.0%)	172 (27.8%)	30.3±36.6
Patients with AAs (n =194)	124	63.8±19.6	75 (38.7%)	27 (13.9%)	53 (27.3%)	73.7±29.0
Control subjects (n = 388)	248	62.6±16.7	83 (21.4%)	43 (11.1%)	119 (30.7%)	8.52±13.2
p value	1.000	0.463	<0.001	0.322	0.405	<0.001

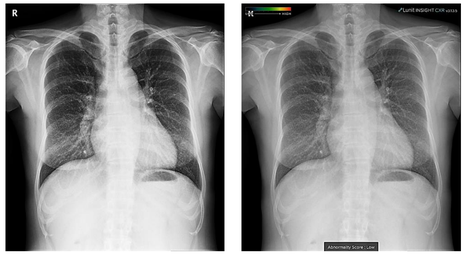
- **Table 2. Diagnostic performance of AI solution and radiologists**

	AI	Reviewer 1	Reviewer 2	Reviewer 3
AUC	95.8%	89.4%	62.0%	83.8%
Sensitivity	90.7%	85.1%	35.6%	70.1%
Specificity	91.2%	93.8%	88.4%	97.4%
Positive rate	39.7%	32.5%	19.6%	25.1%
Detection yield	30.8%	28.4% (p=0.325)	11.9% (p<0.001)	23.4% (p=0.005)
FRR	8.9%	4.1% (p=0.001)	7.7% (p=0.458),	1.7% (p<0.001)

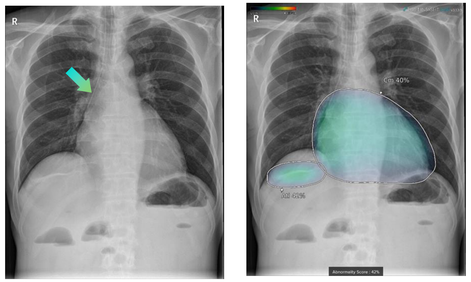
- **Table 3. AI classification of CRs.**

AI classification (Score >22.31)	Total CRs (n = 582)	CRs with AAs (n =194)	Control CRs (n = 388) False Positive
Nodule	54 (25.7%)	50 (28.4%)	4 (11.8%)
Consolidation	47 (22.4%)	43 (24.4%)	4 (11.8%)
Pneumothorax	29 (13.8%)	29 (16.5%)	0 (0.0%)
Pleural effusion	8 (3.8%)	6 (3.4%)	2 (5.9%)
Cardiomegaly	18 (8.6%)	8 (4.5%)	10 (29.4%)
Fibrosis	27 (12.9%)	22 (12.5%)	5 (14.7%)
Pneumoperitoneum	5 (2.4%)	2 (1.1%)	3 (8.8%)
Mediastinal widening	9 (4.3%)	8 (4.5%)	1 (2.9%)
Calcification	4 (1.9%)	1 (0.6%)	3 (8.8%)
Atelectasis	9 (4.3%)	7 (4.0%)	2 (5.9%)
Total	210 (100%)	176 (100%)	34 (100%)

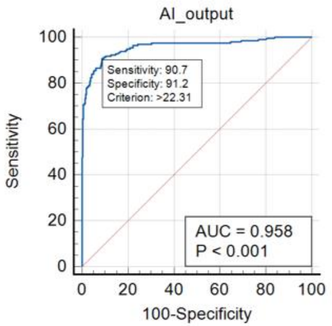
- **Limitations**
  1. A retrospective single center study
  2. Discrepancy from real-world
  - Prevalence of actionable findings
  - Control CRs consisted of normal CRs only
  - Radiologists' decision with limited clinical information



- **Figure 3. A missed actionable finding by AI**: tiny disseminated nodular opacities on both lung zones suggesting miliary tuberculosis were not detected by AI solution but correctly diagnosed by radiologists.



- **Figure 4. A neglected Levin tube by AI**: Levin tube is abnormally located along the airway.



- **Figure 5. ROC curve for the diagnostic performance of AI solution**: An optimal cutoff value of 22.31 yields a sensitivity of 90.7% and a specificity of 91.2% (AUC = 0.958, P < 0.001).

4. CONCLUSION

- **An AI solution could identify actionable abnormalities and enable clinicians to detect critical findings before radiologists' formal interpretation, while its high false referrals should be cautious.**