

Original Article  
Medical Imaging



# Added Value of Bone Suppression Image in the Detection of Subtle Lung Lesions on Chest Radiographs with Regard to Reader's Expertise

Gil-Sun Hong , Kyung-Hyun Do , and Choong Wook Lee

Department of Radiology and Research Institute of Radiology, Asan Medical Center, University of Ulsan College of Medicine, Seoul, Korea



Received: Jun 15, 2019  
Accepted: Aug 19, 2019

**Address for Correspondence:**

**Kyung-Hyun Do, MD, PhD**  
Department of Thoracic Radiology, Asan Medical Center, University of Ulsan College of Medicine, 88 Olympic-ro 43-gil, Songpa-gu, Seoul 05505, Korea.  
E-mail: dokh@amc.seoul.kr

© 2019 The Korean Academy of Medical Sciences.

This is an Open Access article distributed under the terms of the Creative Commons Attribution Non-Commercial License (<https://creativecommons.org/licenses/by-nc/4.0/>) which permits unrestricted non-commercial use, distribution, and reproduction in any medium, provided the original work is properly cited.

**ORCID iDs**

Gil-Sun Hong   
<https://orcid.org/0000-0002-0068-9413>  
Kyung-Hyun Do   
<https://orcid.org/0000-0003-1922-4680>  
Choong Wook Lee   
<https://orcid.org/0000-0001-8776-2603>

**Funding**

This research was supported by a research grant from Samsung Electronics (2017).

**Disclosure**

The authors have no potential conflicts of interest to disclose.

**Author Contributions**

Conceptualization: Hong GS, Do KH. Data curation: Lee CW. Formal analysis: Hong GS.

## ABSTRACT

**Background:** Chest radiographs (CXR) are the most commonly used imaging techniques by various clinicians and radiologists. However, detecting lung lesions on CXR depends largely on the reader's experience level, so there have been several trials to overcome this problem using post-processing of CXR. We investigated the added value of bone suppression image (BSI) in detecting various subtle lung lesions on CXR with regard to reader's expertise.

**Methods:** We applied a software program to generate BSI in 1,600 patients in the emergency department. Of them, 80 patients with subtle lung lesions and 80 patients with negative finding on CXR were retrospectively selected based on the subtlety scores on CXR and CT findings. Ten readers independently rated their confidence in deciding the presence or absence of a lung lesion at each of 960 lung regions on the two separated imaging sessions: CXR alone vs. CXR with BSI.

**Results:** The additional use of BSI for all readers significantly increased the mean area under the curve (AUC) in detecting subtle lung lesions (0.663 vs. 0.706;  $P < 0.001$ ). The less experienced readers were, the more AUC differences increased: 0.067 ( $P < 0.001$ ) for junior radiology residents; 0.064 ( $P < 0.001$ ) for non-radiology clinicians; 0.044 ( $P < 0.001$ ) for senior radiology residents; and 0.019 ( $P = 0.041$ ) for chest radiologists. The additional use of BSI significantly increased the mean confidence regarding the presence or absence of lung lesions for 213 positive lung regions (2.083 vs. 2.357;  $P < 0.001$ ) and for 747 negative regions (1.217 vs. 1.195;  $P = 0.008$ ).

**Conclusion:** The use of BSI increases diagnostic performance and confidence, regardless of reader's expertise, reduces the impact of reader's expertise and can be helpful for less experienced clinicians and residents in the detection of subtle lung lesions.

**Keywords:** Radiography; Computer-Assisted Image Processing; Diagnosis

## INTRODUCTION

Chest radiographs (CXR) are a screening tool for initial surveillance in most cases in the emergency department (ED) and allows for rapid triage in emergency patients. CXR can be also a criterion for determining whether to perform advanced imaging workup and bedside ultrasonography to rapidly identify patients who should start treatment in the ED.<sup>1,2</sup>

Investigation: Hong GS. Methodology: Do KH. Software: Hong GS. Validation: Hong GS, Do KH. Writing - original draft: Hong GS. Writing - review & editing: Do KH, Lee CW.

However, this screening tool may not always be accurate enough, especially in detecting subtle lung lesions in the early stage of diseases such as pneumonia, small lung cancer, or metastasis, because these diseases usually have low contrast and overlying vascular and bony structures can obscure the subtle lung lesions in CXR.<sup>3-8</sup> As a result, normal-looking CXR can delay diagnosis of early stage lung disease in the ED. One previous study suggested that CXR misses nearly half of early stage pneumonia cases associated with influenza.<sup>9</sup> The experience level of clinicians and radiologists also has a great impact on the diagnostic performance of CXR.<sup>10,11</sup> In one prospective multicenter study in patients with suspected community-acquired pneumonia, inter-reader agreement in diagnosing community-acquired pneumonia was only fair to good for experienced radiologists, and poor to fair for inexperienced radiologists and residents.<sup>4,12</sup>

To address these limitations, several techniques such as bone suppression imaging, temporal subtraction techniques, chest tomosynthesis and computer-aided detection have been studied especially for the detection of small lung nodules on chest radiographs.<sup>13-20</sup> Among them, bone suppression image (BSI) has been generated by the dual-energy subtraction (DES) techniques and post-processing soft programs based on machine learning. In recent years, software programs using machine learning have successfully generated bone suppression images.<sup>18,19,21-25</sup> BSI has demonstrated the improved diagnostic performance in detecting focal pneumonia and small lung cancer in the previous studies.<sup>15,16,23,26</sup> One previous study showed that BSI made by dual-energy X-ray equipment can reduce the impact of the reader's expertise in detection of pneumothorax.<sup>27</sup>

In the present study, BSI software from Samsung Electronics (Seoul, Korea) was developed using deep learning; we investigated the added value of BSI in the detection of several subtle lung lesions on CXR that can be frequently seen in the emergency department, regardless of reader's expertise.

## METHODS

This was retrospective case-control study for emergency patients with subtle lung lesions with early stage of the disease.

### Patients

A retrospective review of our institution's database identified 1,600 patients who underwent CXR and Chest computed tomography (CT) within 3 days in the emergency department between October 2016 and March 2017. Based on the Chest CT, 1,291 patients with positive lung lesions were identified. With the reference to CT, two board-certified radiologists, who did not participate in the diagnostic performance study, determined the presence of lung lesions on CXR and rated subtlety of positive lung lesions in consensus on CXR using 6-point scales from no abnormality to vary obvious (**Table 1**). Of 1,291 patients with positive lung lesions, 80 patients who met inclusion criteria were finally enrolled in the present study. The inclusion criteria were: 1) adults over 19 years of age; and 2) patients with early stage of lung disease with a subtlety rating from 1 to 3 on CXR in positive cases. Of 309 patients with negative findings, 80 patients were randomly selected for the control group after age-matching with patients with a positive lung lesion. If exact age-matched patients did not exist, the patients with the smallest age differences were randomly selected.

**Table 1.** Subtlety scoring system of lung lesions on chest radiographs

Scores	Definition
0	Normal chest radiographs despite the presence of lung lesions on chest computed tomography
1	Lesion size $\leq$ 1 cm and/or its density $<$ adjacent vascular structures
2	1 cm $<$ lesion size $\leq$ 2 cm and/or its density similar to adjacent vascular structures
3	2 cm $<$ lesion size $\leq$ 3 cm and/or its density similar to main pulmonary artery
4	Lesion size $>$ 3 cm and/or its density $>$ main pulmonary artery
5	Space occupying lesions replacing one lung field (i.e., huge mass, consolidation or large amount of pleural effusion) on chest radiographs

The definition of early stage disease was based on the assumption that the earlier stage the lung lesions are, the less density and smaller size the lung lesion has on chest radiographs. Therefore, the early stage disease was defined as a lung lesion with a subtlety rating from 1 to 3. For example, early stage pneumonia was defined as infectious community-acquired pneumonia with a subtlety rating from 1 to 3 on chest radiograph, which was characterized by the predominance of peribronchial nodules, ground glass opacities, and/or patchy consolidation on chest CT. The nodular lung cancer was defined as a solitary nodule less than or equal to 3 cm in diameter on chest radiograph, which was a discrete, well-margined nodule with morphological features of lung cancer (i.e., non-calcification, irregular edges, speculated shape, lobulation and/or notching) on chest CT.

### Imaging protocol and generation of bone suppression images

All standard CXRs were performed using a Samsung digital radiography (DR) system (GC85A) (Samsung Electronics; Seoul, Korea) at 120 kVp and 1–2 mAs with individualized automatic exposure control. BSI were generated by Bone Suppression™ software (version 1.02) of the Samsung DR system. This software uses a trained convolutional neural network model to predict bone structures (e.g., ribs and clavicles) on posteroanterior CXRs. The predicted bone structure was then used to construct bone suppression images. This software program was trained with tens of thousands of clinical data of small patch form including normal and abnormal conditions. In this convolutional neural network model, clavicle and rib with anatomical differences were trained separately to improve the ability of deep learning algorithm to recognize and select bony structure on CXR. The sparse kernel method was used to speed up by removing redundant computation during the convolution and pooling operation. Samsung Electronics provided the software, but they did not have any influence over the image selection and observer study design.

### Image analysis

All images were reviewed using a local picture archiving and communication system (PACS) monitor and digital imaging and communications in medicine image-viewing software (PetaVision; Asan Medical Center, Seoul, Korea). Two board-certified radiologists rated subtlety of positive lung lesions in consensus on CXR using 6-point scales from no abnormality to vary obvious (Table 1). Even if positive lung lesions, such as ground-glass opacity nodules, were present on chest CT, they were rated a subtlety score of 0 and excluded from the study if no lesion could be found on CXR by two board-certified radiologists.

Ten readers (three board-certified chest radiologists, two board-certified non-radiology clinicians, three senior radiology residents and two junior radiology residents), who were blind to the patient diagnoses and clinical information, independently evaluated 160 cases in a visual manner in two separated reading sessions with a two-week interval in order to avoid recall bias. In each reading session, all 160 cases were randomly re-ordered. Lung

field of CXR was divided into six regions (upper, middle and lower zones in both right and left lungs) on the basis of the anatomical structures on CXR. For a total of 960 regions, ten readers rated their confidence for the presence or absence of lung lesion in each region using a 5-grade scale: 1, definite normal; 2, probable normal; 3, indeterminate (clinical observation is recommended); 4, probable positive lesion; and 5, definite positive lesion. In the first reading session, CXRs alone were displayed in a full-scale and interpreted by the readers. In the second reading session, two chest radiographs (CXR with BSI) were displayed in a full-scale and side by side and interpreted by the readers. A single investigator measured the size of small lung nodules and nodular lung cancers on chest radiographs using PACS software. In the case of multiple metastases, the size of the largest nodule in each region was measured.

### Statistical analysis

The characteristics of patients between two groups were compared using independent *t*-test and Pearson's  $\chi^2$  test. Confidence levels of grade 3–5 were considered as deciding the presence of a positive lung lesion at each region. The statistical significance of the difference in the values of the area under the receiver operating characteristic (ROC) curve (AUC) between two reading sessions (CXR vs. CXR with BSI) was tested by use of the Dorfman-Berbaum-Metz method,<sup>28,29</sup> which included multireader, multicase and multilocation variation by means of an analysis of variance approach (DBM MRMC, version 2.2; <http://rocweb.bsd.uchicago.edu/MetzROC>). The AUC values for each subgroup with respect to the subtlety scores were calculated and compared using Obuchowski method, in which the control groups with negative chest radiographs were used equally when calculating the AUC values. General estimation equations were used for comparisons of per-region detection rate to take multiple lung lesions per regions into account between two reading sessions. To assess the effect of BSI on reader confidence in diagnosing the presence or absence of pulmonary lesions, we compared the confidence grade between two reading sessions using Wilcoxon signed-rank test. The complete data on diagnostic confidence was presented on heat maps.<sup>30</sup> SAS version 9.4 (SAS Institute, Cary, NC, USA) & R 3.4.2 was used to perform statistical comparisons and heat maps. A two-tailed parameter with a significance threshold of  $P < 0.05$  was used.

### Ethics statement

The present study protocol was reviewed and approved by the Institutional Review Board of Asan Medical Center (approval No. 2017-0197). Written informed consent to the use of CXR and CT images for research and publication was waived by the Institutional Review Board due to retrospective study design.

## RESULTS

### Clinical characteristics of the study population

**Table 2** summarizes the baseline demographics and clinical characteristics of the study patients. The present study consisted of patients with early stage pneumonia ( $n = 41$  of 80 patients, 51.3%), multiple metastasis ( $n = 19$ , 23.8%), small solitary lung nodule ( $n = 5$ , 6.3%), nodular lung cancer ( $n = 9$ , 11.3%), and non-tuberculous mycobacterial infection or pulmonary tuberculosis (NTM/TB) ( $n = 6$ , 7.5%). **Table 2** also shows the detailed number of positive lung regions with respect to the location and subtlety scores. The mean sizes  $\pm$  standard deviation (SD) of solitary lung nodules and nodular lung cancers were  $9.4 \pm 3.9$  mm and  $24.0 \pm 4.6$  mm, respectively. The mean sizes  $\pm$  SD of metastatic nodules were  $8.0 \pm 3.9$  mm.

**Table 2.** Characteristics of all study patients with positive and negative lung lesions

Characteristics	Patients with positive lung lesions (n = 80)	Patients with negative lung lesions (n = 80)	P value
Age, yr	62.8 ± 14.3	61.2 ± 14.3	0.933
Gender, men:women	49:31	43:37	0.008
Time interval between CXR and CT, day <sup>a</sup>	0.3 ± 0.6 (0–3)	0.2 ± 0.4 (0–1)	0.723
Diagnosis on CT <sup>b</sup>			
Early stage pneumonia	41 (120)	N/A	
Metastasis	19 (66)	N/A	
Small lung nodules	5 (5)	N/A	
Nodular lung cancer	9 (9)	N/A	
NTM/TB	6 (13)	N/A	
Location of lung lesions <sup>c</sup>			
RULZ	36	N/A	
RMLZ	37	N/A	
RLLZ	56	N/A	
LULZ	20	N/A	
LMLZ	25	N/A	
LLLZ	39	N/A	
Subtlety scores <sup>c</sup>			
1	85	N/A	
2	75	N/A	
3	53	N/A	

CXR = standard chest radiographs, CT = computed tomography, NTM = nontuberculous mycobacteria, TB = pulmonary tuberculosis, RULZ = right upper lung zone, RMLZ = right middle lung zone, RLLZ = right lower lung zone, LULZ = left upper lung zone, LMLZ = left middle lung zone, LLLZ = left lower lung zone.

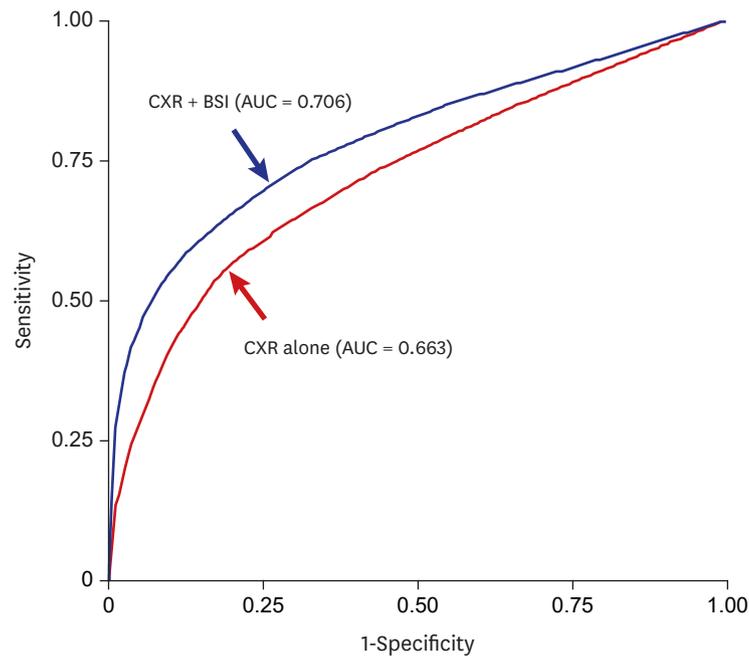
<sup>a</sup>Data are shown as mean ± standard deviation with the range of data in parentheses; <sup>b</sup>Data are shown as number of patients with the number of lung regions with positive lung lesion in parentheses; <sup>c</sup>Data are shown as number of lung regions with positive lung lesion.

### Diagnostic performance

The mean AUC value per region for all readers significantly increased in detecting subtle lung lesions between two sessions (CXR alone vs. CXR with BSI; 0.663 vs. 0.706;  $P < 0.001$ ) (Fig. 1 and Table 3). The less experienced the readers were, the more AUC differences between the two sessions increased: the mean AUC differences were 0.067 ( $P < 0.001$ ) for junior radiology residents; 0.064 ( $P < 0.001$ ) for non-radiology clinicians; 0.044 ( $P < 0.001$ ) for senior radiology residents; and 0.019 ( $P = 0.041$ ) for chest radiologists. The per-regional detection rates of each disease are summarized in Fig. 2. The per-regional detection rates were significantly increased in early stage pneumonia (36.3% vs. 43.3%;  $P < 0.001$ ), metastasis (31.7% vs. 41.8%;  $P < 0.001$ ), and NTM/TB (46.9% vs. 57.7%;  $P = 0.004$ ) between two sessions (Figs. 3 and 4). With BSI, the per-regional detection rates of small lung nodules (32.0% vs. 46.0%;  $P = 0.052$ ) and nodular lung cancer (55.6% vs. 62.2%;  $P = 0.083$ ) slightly increased but not significantly. Table 4 summarizes the results of subgroup analysis of AUC value with respect to the location and subtlety scores of lung lesion: the mean AUC value significantly increased between two sessions (CXR alone vs. CXR with BSI) regardless of the location of lung lesion (all  $P < 0.05$ ); and the AUC difference between two sessions was the highest in the subtlety score = 1 (0.054), followed by the subtlety score = 2 (0.037) and the subtlety score = 3 (0.036). In the present study, there was no significant difference in the mean false positive rates for 10 readers between two sessions (CXR alone vs. CXR with BSI; 6.6% [95% CI, 6.1%–7.2%] vs. 6.4% [95% CI, 5.8%–6.9%];  $P = 0.376$ ).

### Diagnostic confidence

The confidence level regarding the presence or absence of subtle lung lesions are summarized in Table 5. The mean confidence regarding the presence of subtle lung lesions significantly improved between session 1 and 2 for 213 positive lung regions (2.083 vs.



**Fig. 1.** The mean ROC curves for per-regional detection of subtle lung lesions on CXR without and with BSI by ten readers. AUC = area under the ROC curve, ROC = receiver operating characteristic, CXR = standard chest radiographs, BSI = bone suppression image.

**Table 3.** Per-regional detection of subtle lung lesions on CXRs without or with BSI for ten readers

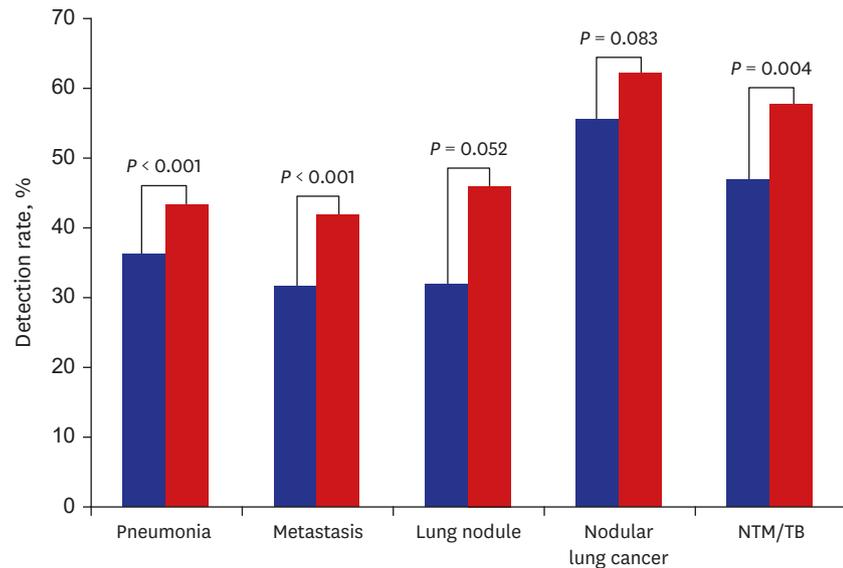
Variables	CXR AUC	CXR + BSI AUC	AUC difference (95% CI)	P value
Total	0.663	0.706	0.043 (0.032, 0.055)	< 0.001
Group 1				
Reader 1	0.681	0.683	0.002 (-0.033, 0.037)	0.912
Reader 2	0.768	0.799	0.030 (-0.001, 0.062)	0.061
Reader 3	0.722	0.747	0.025 (-0.010, 0.060)	0.163
Mean	0.722	0.741	0.019 (0.001, 0.037)	0.041
Group 2				
Reader 4	0.693	0.744	0.050 (0.010, 0.091)	0.014
Reader 5	0.521	0.602	0.081 (0.044, 0.118)	< 0.001
Mean	0.606	0.670	0.064 (0.040, 0.088)	< 0.001
Group 3				
Reader 6	0.666	0.714	0.048 (-0.004, 0.010)	0.069
Reader 7	0.683	0.716	0.033 (-0.021, 0.087)	0.230
Reader 8	0.728	0.782	0.054 (0.009, 0.099)	0.019
Mean	0.690	0.733	0.044 (0.014, 0.073)	0.004
Group 4				
Reader 9	0.603	0.669	0.066 (0.036, 0.096)	< 0.001
Reader 10	0.566	0.635	0.069 (0.035, 0.103)	< 0.001
Mean	0.585	0.652	0.067 (0.047, 0.088)	< 0.001

Data are AUCs values.

CXR = standard chest radiographs, BSI = bone suppression image, AUC = area under the ROC curve, CI = confidence interval.

Group 1 = chest radiologists; Group 2 = non-radiology clinicians; Group 3 = senior radiology residents; and Group 4 = junior radiology residents.

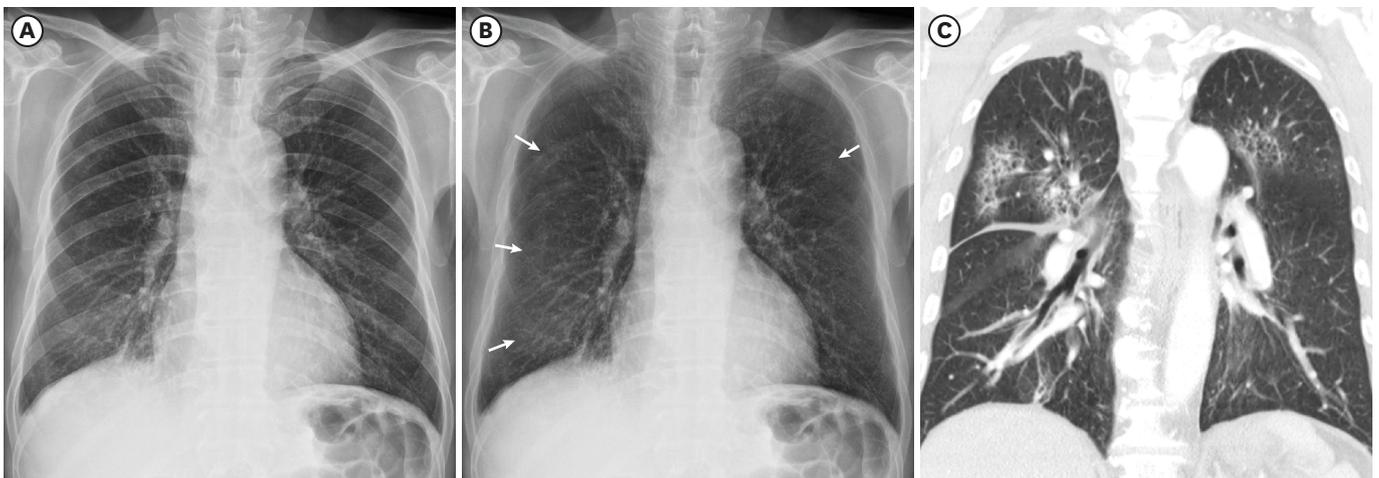
2.357;  $P < 0.001$ ) and for 747 negative regions (1.217 vs. 1.195;  $P = 0.008$ ). All data regarding diagnostic confidence are illustrated in heat maps with clustering (Fig. 5). For positive lung regions, higher confidence grades (i.e., red color) increased and indeterminate or lower confidence grades (i.e., orange or yellow color) decreased in positive lung regions when



**Fig. 2.** The per-regional detection rate for five lung diseases on CXR without or with BSI.

CXR = standard chest radiographs, BSI = bone suppression image, NTM/TB = non-tuberculous mycobacterial infection or pulmonary tuberculosis.

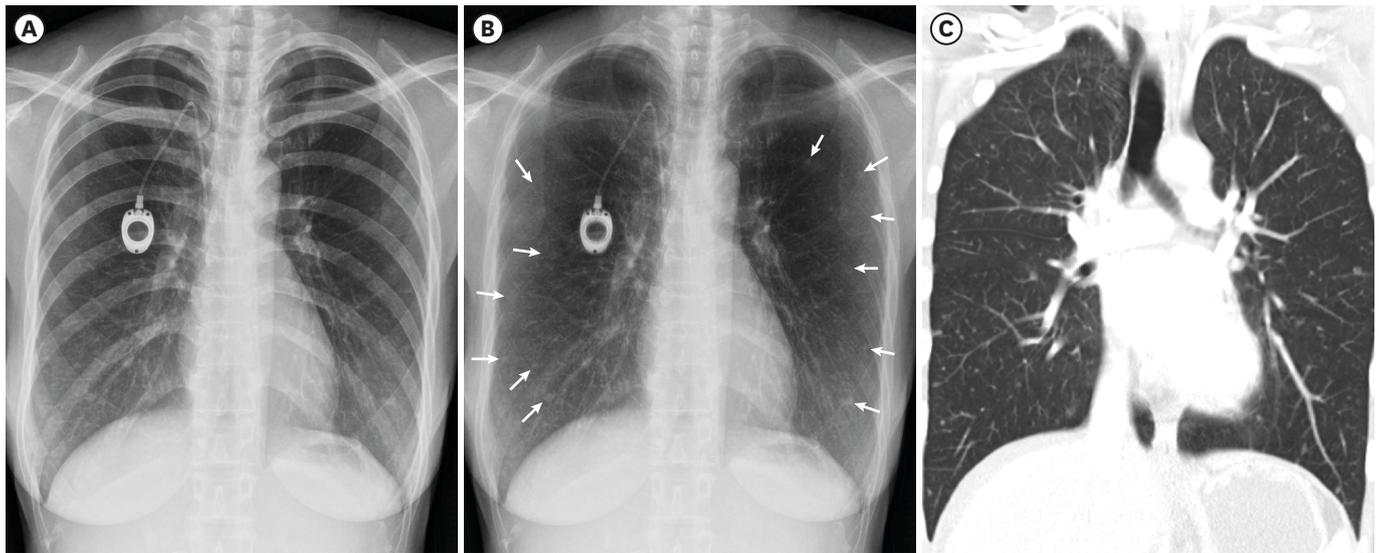
adding BSI to CXR. For negative lung regions, heat map also showed some color changes. The clustered difference map illustrated the switching of readers' confidence according to the use of additional BSI. Readers converted the confidence rating in the direction of increased confidence (i.e., blue color) in the positive lung regions, while in the negative lung regions, the pattern of these changes was not noticeable. In **Fig. 5A and B**, the dendrogram along the x-axis showed that there was major separation between two groups of readers and by looking at the colors, one can see that one group (Group A) represented the high degree of readers' confidence and another group (Group B) represented the relatively lower degree of readers' confidence in the positive lung regions. **Fig. 5A** showed that each group consisted of the same number of readers when using CXR alone and the heights of the arms of the dendrograms at each group are relatively different. However, when using CXR with BSI (**Fig. 5B**), as the



**Fig. 3.** A 62-year-old male with atypical pneumonia confirmed by CT. **(A)** CXR shows patch increased opacity in right upper lung area. **(B)** BSI shows multiple increased opacities in bilateral lungs more clearly. **(C)** CT shows multiple ground glass opacities and nodules in bilateral lungs.

CT = computed tomography, CXR = standard chest radiographs, BSI = bone suppression image.

Written informed consent to the use of CT images for research and publication was waived by the Institutional Review Board due to retrospective study design.



**Fig. 4.** A 43-year-old female with multiple metastases in bilateral lungs confirmed on CT. **(A)** CXR shows nodular opacities in the left upper and middle lung areas. **(B)** BSI shows multiple nodular nodules more clearly in bilateral lungs. **(C)** CT shows multiple metastatic nodules in bilateral lungs. CT = computed tomography, CXR = standard chest radiographs, BSI = bone suppression image. Written informed consent to the use of CT images for research and publication was waived by the Institutional Review Board due to retrospective study design.

**Table 4.** Subgroup analysis of mean AUC value with respect to the location and subtlety scores of lung lesion

Variables	CXR AUC	CXR + BSI AUC	AUC difference (95% CI)	P value
<b>Location<sup>a</sup></b>				
RULZ	0.669	0.700	0.031 (0.055, 0.006)	0.014
RMLZ	0.676	0.714	0.038 (0.062, 0.014)	0.002
RLLZ	0.678	0.742	0.064 (0.087, 0.041)	< 0.001
LULZ	0.638	0.680	0.042 (0.079, 0.004)	0.030
LMLZ	0.625	0.675	0.051 (0.080, 0.021)	0.001
LLLZ	0.628	0.676	0.049 (0.077, 0.021)	0.001
<b>Subtlety score<sup>b</sup></b>				
1	0.577	0.631	0.054 (0.037, 0.072)	< 0.001
2	0.675	0.711	0.037 (0.017, 0.056)	< 0.001
3	0.782	0.818	0.036 (0.015, 0.056)	0.001

AUC = area under the ROC curve, CXR = standard chest radiographs, BSI = bone suppression image, CI = confidence interval, RULZ = right upper lung zone, RMLZ = right middle lung zone, RLLZ = right lower lung zone, LULZ = left upper lung zone, LMLZ = left middle lung zone, LLLZ = left lower lung zone.

<sup>a</sup>Data are AUCs values, calculated using Dorfman-Berbaum-Metz significance testing, with Hillis improvements;

<sup>b</sup>Data are AUCs values, calculated using Obuchowski method.

readers' confidence increased, most of readers were grouped into one large group (Group B) except for 3 readers who had the highest confidence and the difference in the height of the arm of the dendrograms between readers was reduced. In the negative lung regions, the dendrograms showed a similar pattern to the positive lung regions, but the difference in the heights of the arms of the dendrograms between readers was slightly reduced when using CXR with BSI.

## DISCUSSION

In the present study, we evaluated the diagnostic performance of readers with different levels of experience in detecting subtle lung lesion on CXR when using additional BSI. The level of experience of readers is one main factor affecting the results in the observer

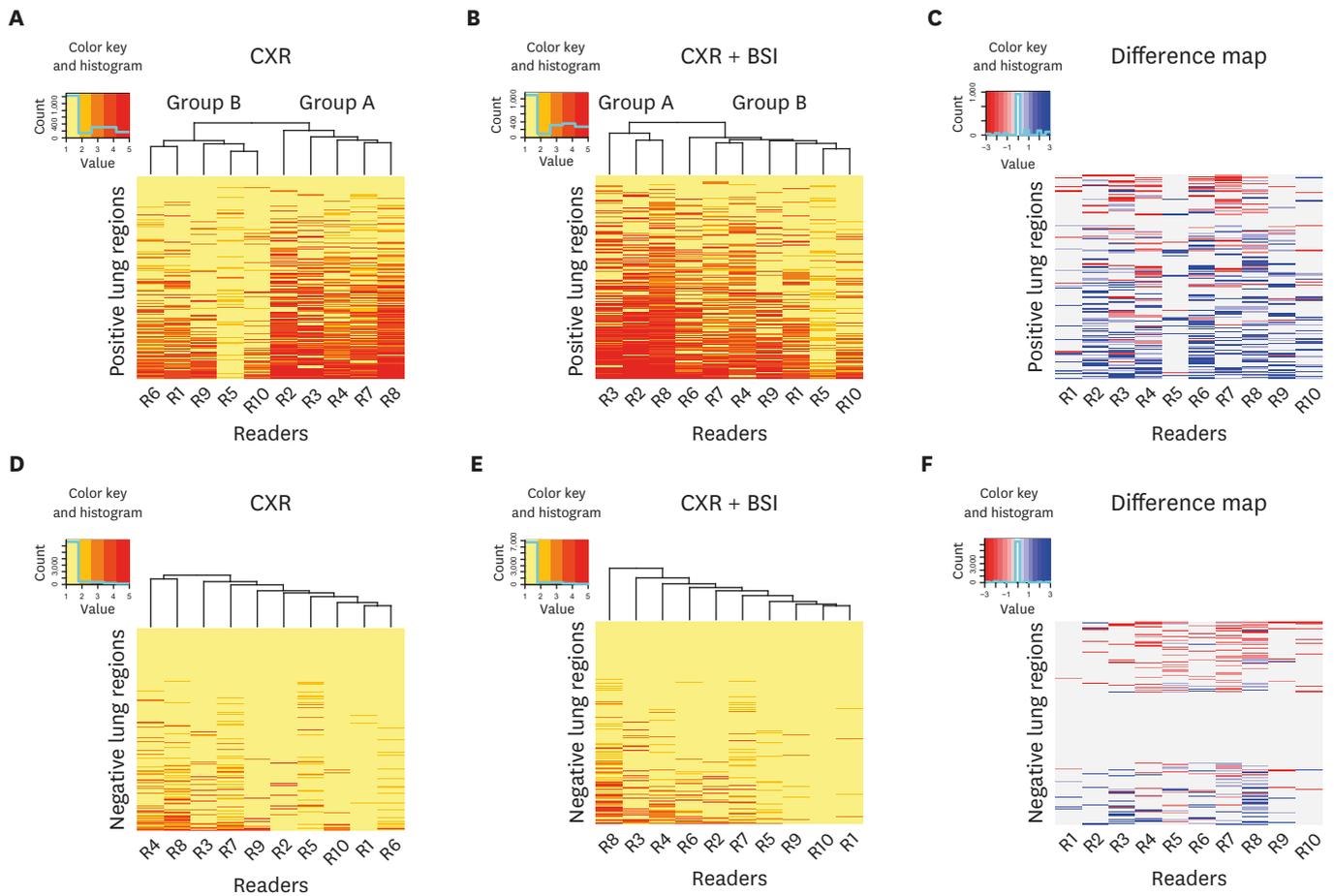
**Table 5.** Per-regional confidence level regarding the presence of subtle lung lesions on CXR without or with BSI for ten observers

Groups	Positive lung regions (n = 213)			Negative lung regions (n = 747)		
	CXR	CXR + BSI	P value	CXR	CXR + BSI	P value
Total	2.083	2.357	< 0.001	1.217	1.195	0.008
Group 1						
Reader 1	1.911	1.991		1.040	1.055	
Reader 2	2.761	3.188		1.126	1.137	
Reader 3	2.624	2.901		1.250	1.299	
Mean	2.432	2.693	< 0.001	1.139	1.163	0.069
Group 2						
Reader 4	2.366	2.455		1.414	1.277	
Reader 5	1.211	1.358		1.181	1.080	
Mean	1.789	1.908	0.008	1.297	1.179	< 0.001
Group 3						
Reader 6	1.850	2.394		1.104	1.189	
Reader 7	2.277	2.286		1.348	1.186	
Reader 8	2.732	3.216		1.454	1.640	
Mean	2.286	2.632	< 0.001	1.302	1.338	0.027
Group 4						
Reader 9	1.690	2.099		1.142	1.052	
Reader 10	1.404	1.681		1.108	1.035	
Mean	1.547	1.890	< 0.001	1.125	1.044	< 0.001

Data are mean confidence scores using 5 scales (1: definite normal; 2: probable normal; 3: indeterminate; 4: probable positive lesion; and 5: definite positive lesion). Group 1 = chest radiologists; Group 2 = non-radiology clinicians; Group 3 = senior radiology residents; and Group 4 = junior radiology residents. CXR = standard chest radiographs, BSI = bone suppression image.

test.<sup>31-33</sup> Li et al.<sup>23</sup> reported that regardless of readers, the diagnostic performance in the detection of lung nodule was significantly improved by the use of BSI. In the present study, the use of BSI significantly improved the pre-regional diagnostic performance of all reader groups in the detection of subtle lung lesions. The non-radiology clinician, junior and senior radiology resident groups were inferior to chest radiologists in detecting subtle lung lesions using CXR alone. The additional use of BSI was particularly effective in increasing the diagnostic performance of non-radiology clinicians and radiology resident groups, and consequently reduced the inter-reader difference in identifying positive lung regions. In the senior radiology resident group, the per-regional diagnostic performance was improved close to that of chest radiologist group when using the additional BSI. These results could be worthwhile in that CXR is an imaging test used by various clinicians with different experience levels in the ED. Therefore, our study indicates that the addition of BSI to CXR could provide a good method to reduce the impact of the clinicians' experience level, which can be helpful for clinicians with less experience and residents to rapidly identify patients with subtle lung lesions and early stage disease who should start treatment in the ED.

Our current study showed that the addition of BSI to CXR increased diagnostic confidence and decreases confidence variability among readers when interpreting positive lung lesions. The heat maps shown in the present study illustrated that the addition of BSI to CXR increased red color in the positive lung regions, whereas BSI decreased yellow color (i.e., less definite decision) or orange color (i.e., indeterminate decision) in positive lung regions. The clustered difference map summarized this effect of the use of additional BSI on the readers' confidence level at a glance. This result suggested that the use of BSI shifts the readers' interpretation of the positive lung regions toward a high level of confidence. The dendrograms pattern represents a hierarchy of associations where the mathematical distance between groups is illustrated as the short or long arm (i.e., high association between groups or relatively distant association between groups). Therefore, the grouping on the dendrogram does not indicate that the readers were correct but that they agree with each other. In the



**Fig. 5.** Heat maps with clustering and dendrograms of per-regional diagnostic confidences in the 213 positive and 747 negative lung regions as interpreted by 10 readers. **(A)** Confidence rating on the positive lung regions with CXR alone. **(B)** Confidence rating on the positive lung regions with CXR with BSI. **(C)** the clustered difference map on the positive lung regions. **(D)** Confidence rating on the negative lung regions with CXR alone. **(E)** Confidence rating on the negative lung regions with CXR with BSI. **(F)** the clustered difference map on the negative lung regions. On the heat maps, each grid element at the intersection of rows and columns represents the degree of diagnostic confidence of each reader (column) for each corresponding region (row). From yellow to red, each color corresponds to a confidence level of 1 to 5 levels. In the case of positive lungs, the lattice elements become increasingly confident as they move from yellow to red color. In negative lungs, confidence increases in the opposite direction. Orange color means an indeterminate decision. To show the distribution of colors at a glance, the difference maps are fixed along the x axis and clustered along the y axis. At the clustered difference maps, blue color indicates that the readers' confidence is switched toward the high level in both the positive and negative lung regions, whereas red color indicates the opposite direction. CXR = standard chest radiographs, BSI = bone suppression image.

present study, using additional BSI grouped most readers into one large group and reduced the difference in the height of the arm of the dendrograms between readers in positive lung regions, which means that the readers with different experience levels had more similarity (i.e., the increased association between readers). In our study, the use of BSI increased the mean confidence level in the negative lung regions but had a small effect on the distance of the association between readers. There was no difference in the mean false positive rates between two reading sessions (BSI alone vs. CXR with BSI) in the present study. However, the use of BSI has the potential risk of causing misinterpretation of normal structures as abnormal lesions because bone subtraction increases the prominence of focal scars and normal vascular structures. We think that this issue seems not to be resolved considering the difficulty in calculating the potential risk of the misinterpretation because it is influenced by the reader's expertise and it takes a long time for the reader to know the limitations of the new technology and to distinguish artificial findings from real lesions. In summary, an overall increase in diagnostic confidence and a reduction in the variability of the readers'

confidence may be important for clinicians and residents in rapidly deciding whether to perform additional work-up or start treatment in ED patients with subtle lung lesions.

Our study included five types of subtle lung lesions commonly encountered in the ED. Pneumonia with faint focal or patchy opacities can be easily missed on CXR, even for experienced radiologists and non-radiology clinicians. Ojutiku et al.<sup>7</sup> reported the only 29 of 60 pneumonia (48%) was detected by radiology residents. Consistent with the previous study,<sup>20</sup> our study showed that the detection rate of early stage pneumonia significantly increased by the use of BSI. Unlike the previous study with focal nodular pneumonia mimicking a nodular lung cancer,<sup>20</sup> our study included 12 patients with atypical pneumonia seen as multifocal ground-glass opacity on chest CT. Therefore, it is likely to be effective in the diagnosis of various diseases (e.g., lung contusion and blood aspiration) with CT findings similar to atypical pneumonia. Our study also showed that the detection rates of metastatic nodules and NTM/TB significantly increased by the use of BSI. However, unlike previous studies,<sup>15,16,19,22-24,26</sup> the per-regional detection rates of small nodules and nodular lung cancers in our study did not increase with the use of BSI, probably due to the insufficient number of cases. This explanation is reasonable considering that the use of BSI significantly increased the per-regional detection rate of metastatic nodules, although metastatic nodules are generally smaller than nodular lung cancer.

In the present study, we anticipated that using BSI could be useful in detecting early stage lung lesions and assumed that early stage lung lesions generally manifest as subtle lung lesions (low density and/or small size of lesions) on the chest radiographs. The interesting finding in the subgroup analysis was that the efficacy of BSI varied depending on the subtlety scores. As our expectation, the additional use of BSI had much greater impact on the detection of lung lesions with a lower subtlety score more than the detection of lesions with a higher subtlety score. Bone suppression images are expected to be most useful in detecting the lesions in bilateral upper lung fields because they can be easily obscured by the clavicles and upper ribs. It is well-known that missed lung cancer is mostly located in the upper lobes.<sup>34</sup> However, in the present study, there was no difference in detecting lung lesions according to the location of the lesions. This may be due to the lack of a sufficient number of suitable cases (i.e., lung lesions in upper lung fields obscured by the bony structures) to prove this assumption in our study. In our study, pneumonia accounted for the largest portion of the study group, followed by metastasis. Pneumonia can occur anywhere but is sometimes gravity dependent and depends on the person's position. The metastasis has peripheral distribution and lower lung predominance due to the prolonged transit time in the peripheral lung and the relative basilar overperfusion. Lung cancer accounted for a relatively small portion of the present study. Therefore, further studies that include enough suitable cases located in the upper lung fields are necessary to improve our assumption.

To date, dual energy subtraction techniques (DES) have been a standard method to generate bone suppression images.<sup>15,16,23,26</sup> However, DES technique has some disadvantages like a small potential increase in radiation dose, misregistration artifact due to respiration motion and low signal-to-noise ratio. In the case of post-processing software based on machine learning, the technical efficacy of programs in generating BSI in various clinical settings remains unclear, even though it has provided similar images to BSI generated by DES techniques. The deep learning-based software program of this study is expected to increase the efficiency, reproducibility, and coverage of the software in generating BSI and be more applicable to the various clinical settings than the pre-existing software programs. We think

that the efficacy of these software programs in generating BSI in the various clinical settings is an important issue that requires further research.

Our study had several limitations of note. Its retrospective design may have introduced selection bias, as the study population consisted of selected cases under controlled conditions. Our study did not include all subtle lung lesions encountered in the emergency department because the observer performance study is affected by the limited diagnostic ability of imaging modality (i.e. chest radiographs) compared with that of chest CT. However, this is an inevitable bias in the observer performance study using the imaging modality with limited diagnostic ability. The readers' performance of BSI largely depends on the prevalence of bone-overlapped subtle lung lesions. Our case-control study cannot reflect a real-world prevalence of these lung lesions in emergency patients. In addition, we did not evaluate the efficiency and coverage of the commercial software to successfully generate BSI in various clinical settings. Therefore, the general application of this technique in the ED needs further validation at other institutions. Second, the sequential presentation of images in our study may have introduced bias based on the assumption that the CXR plus BSI would be superior to CXR alone. However, the sequential presentation of images should not be a problem since this type of previous research had no significant benefit in detection accuracy.<sup>23</sup> In addition, the lungs were divided into six regions and each lesion was evaluated. This research design may have introduced bias in that the evaluation of one region may affect the assessment of other regions. However, Li's study<sup>23</sup> showed consistent results when independent scoring of each lung, even if some degree of correlation between the pairs of lungs exists.

In conclusion, we showed that the use of BSI increases diagnostic performance and confidence, regardless of reader's expertise. It also reduces the impact of reader's expertise and can be helpful for less experienced clinicians and residents in the detection of subtle lung lesions that can be frequently seen in the emergency department.

## ACKNOWLEDGMENTS

The authors are grateful to Han Na Noh, MD, Hyun Joo Lee, MD, Jooae Choe, MD, Dong Kyu Oh, MD, Hyung-Joo Lee, MD, Hyo-Jung Park, MD, Ji Hye Kwon, MD, Ho Young Park, MD, Su Jin Lim, MD, and Sohee Park, MD for participating as readers and acknowledge Seon-Ok Kim in the department of clinical epidemiology and biostatistics, Asan Medical Center for statistical advice.

## REFERENCES

1. Rodriguez RM, Anglin D, Langdorf MI, Baumann BM, Hendey GW, Bradley RN, et al. NEXUS chest: validation of a decision instrument for selective chest imaging in blunt trauma. *JAMA Surg* 2013;148(10):940-6.  
[PUBMED](#) | [CROSSREF](#)
2. Rodriguez RM, Langdorf MI, Nishijima D, Baumann BM, Hendey GW, Medak AJ, et al. Derivation and validation of two decision instruments for selective chest CT in blunt trauma: a multicenter prospective observational study (NEXUS Chest CT). *PLoS Med* 2015;12(10):e1001883.  
[PUBMED](#) | [CROSSREF](#)
3. Young M, Marrie TJ. Interobserver variability in the interpretation of chest roentgenograms of patients with possible pneumonia. *Arch Intern Med* 1994;154(23):2729-32.  
[PUBMED](#) | [CROSSREF](#)

4. Albaum MN, Hill LC, Murphy M, Li YH, Fuhrman CR, Britton CA, et al. Interobserver reliability of the chest radiograph in community-acquired pneumonia. *Chest* 1996;110(2):343-50.  
[PUBMED](#) | [CROSSREF](#)
5. Ketai L, Malby M, Jordan K, Meholic A, Locken J. Small nodules detected on chest radiography: does size predict calcification? *Chest* 2000;118(3):610-4.  
[PUBMED](#) | [CROSSREF](#)
6. Shah PK, Austin JH, White CS, Patel P, Haramati LB, Pearson GD, et al. Missed non-small cell lung cancer: radiographic findings of potentially resectable lesions evident only in retrospect. *Radiology* 2003;226(1):235-41.  
[PUBMED](#) | [CROSSREF](#)
7. Ojutiku O, Haramati LB, Rakoff S, Sprayregen S. Radiology residents' on-call interpretation of chest radiographs for pneumonia. *Acad Radiol* 2005;12(5):658-64.  
[PUBMED](#) | [CROSSREF](#)
8. Gourin CG, Watts TL, Williams HT, Patel VS, Bilodeau PA, Coleman TA. Identification of distant metastases with positron-emission tomography-computed tomography in patients with previously untreated head and neck cancer. *Laryngoscope* 2008;118(4):671-5.  
[PUBMED](#) | [CROSSREF](#)
9. Testa A, Soldati G, Copetti R, Giannuzzi R, Portale G, Gentiloni-Silveri N. Early recognition of the 2009 pandemic influenza A (H1N1) pneumonia by chest ultrasound. *Crit Care* 2012;16(1):R30.  
[PUBMED](#) | [CROSSREF](#)
10. Kelly BS, Rainford LA, Darcy SP, Kavanagh EC, Toomey RJ. The development of expertise in radiology: in chest radiograph interpretation, "expert" search pattern may predate "expert" levels of diagnostic accuracy for pneumothorax identification. *Radiology* 2016;280(1):252-60.  
[PUBMED](#) | [CROSSREF](#)
11. Mehrotra P, Bosemani V, Cox J. Do radiologists still need to report chest x rays? *Postgrad Med J* 2009;85(1005):339-41.  
[PUBMED](#) | [CROSSREF](#)
12. Melbye H, Dale K. Interobserver variability in the radiographic diagnosis of adult outpatient pneumonia. *Acta Radiol* 1992;33(1):79-81.  
[PUBMED](#)
13. McAdams HP, Samei E, Dobbins J 3rd, Tourassi GD, Ravin CE. Recent advances in chest radiography. *Radiology* 2006;241(3):663-83.  
[PUBMED](#) | [CROSSREF](#)
14. Girvin F, Ko JP. Pulmonary nodules: detection, assessment, and CAD. *AJR Am J Roentgenol* 2008;191(4):1057-69.  
[PUBMED](#) | [CROSSREF](#)
15. Li F, Engelmann R, Doi K, MacMahon H. Improved detection of small lung cancers with dual-energy subtraction chest radiography. *AJR Am J Roentgenol* 2008;190(4):886-91.  
[PUBMED](#) | [CROSSREF](#)
16. MacMahon H, Li F, Engelmann R, Roberts R, Armato S. Dual energy subtraction and temporal subtraction chest radiography. *J Thorac Imaging* 2008;23(2):77-85.  
[PUBMED](#) | [CROSSREF](#)
17. Dobbins JT 3rd, McAdams HP. Chest tomosynthesis: technical principles and clinical update. *Eur J Radiol* 2009;72(2):244-51.  
[PUBMED](#) | [CROSSREF](#)
18. Oda S, Awai K, Suzuki K, Yanaga Y, Funama Y, MacMahon H, et al. Performance of radiologists in detection of small pulmonary nodules on chest radiographs: effect of rib suppression with a massive-training artificial neural network. *AJR Am J Roentgenol* 2009;193(5):W397-402.  
[PUBMED](#) | [CROSSREF](#)
19. Li F, Hara T, Shiraishi J, Engelmann R, MacMahon H, Doi K. Improved detection of subtle lung nodules by use of chest radiographs with bone suppression imaging: receiver operating characteristic analysis with and without localization. *AJR Am J Roentgenol* 2011;196(5):W535-41.  
[PUBMED](#) | [CROSSREF](#)
20. Li F, Engelmann R, Pesce L, Armato SG 3rd, Macmahon H. Improved detection of focal pneumonia by chest radiography with bone suppression imaging. *Eur Radiol* 2012;22(12):2729-35.  
[PUBMED](#) | [CROSSREF](#)
21. Suzuki K, Abe H, MacMahon H, Doi K. Image-processing technique for suppressing ribs in chest radiographs by means of massive training artificial neural network (MTANN). *IEEE Trans Med Imaging* 2006;25(4):406-16.  
[PUBMED](#) | [CROSSREF](#)

22. Freedman MT, Lo SC, Seibel JC, Bromley CM. Lung nodules: improved detection with software that suppresses the rib and clavicle on chest radiographs. *Radiology* 2011;260(1):265-73.  
[PUBMED](#) | [CROSSREF](#)
23. Li F, Engelmann R, Pesce LL, Doi K, Metz CE, Macmahon H. Small lung cancers: improved detection by use of bone suppression imaging--comparison with dual-energy subtraction chest radiography. *Radiology* 2011;261(3):937-49.  
[PUBMED](#) | [CROSSREF](#)
24. Schalekamp S, van Ginneken B, Meiss L, Peters-Bax L, Quekel LG, Snoeren MM, et al. Bone suppressed images improve radiologists' detection performance for pulmonary nodules in chest radiographs. *Eur J Radiol* 2013;82(12):2399-405.  
[PUBMED](#) | [CROSSREF](#)
25. Miyoshi T, Yoshida J, Aramaki N, Matsumura Y, Aokage K, Hishida T, et al. Effectiveness of bone suppression imaging in the detection of lung nodules on chest radiographs: relevance to anatomic location and observer's experience. *J Thorac Imaging* 2017;32(6):398-405.  
[PUBMED](#) | [CROSSREF](#)
26. Szucs-Farkas Z, Patak MA, Yuksel-Hatz S, Ruder T, Vock P. Single-exposure dual-energy subtraction chest radiography: detection of pulmonary nodules and masses in clinical practice. *Eur Radiol* 2008;18(1):24-31.  
[PUBMED](#) | [CROSSREF](#)
27. Urbaneja A, Dodin G, Hossu G, Bakour O, Kechidi R, Gondim Teixeira P, et al. Added value of bone subtraction in dual-energy digital radiography in the detection of pneumothorax: impact of reader expertise and medical specialty. *Acad Radiol* 2018;25(1):82-7.  
[PUBMED](#) | [CROSSREF](#)
28. Dorfman DD, Berbaum KS, Metz CE. Receiver operating characteristic rating analysis. Generalization to the population of readers and patients with the jackknife method. *Invest Radiol* 1992;27(9):723-31.  
[PUBMED](#) | [CROSSREF](#)
29. Hillis SL, Berbaum KS, Metz CE. Recent developments in the Dorfman-Berbaum-Metz procedure for multireader ROC study analysis. *Acad Radiol* 2008;15(5):647-61.  
[PUBMED](#) | [CROSSREF](#)
30. Freedman MT, Osicka T. Heat maps: an aid for data analysis and understanding of ROC CAD experiments. *Acad Radiol* 2008;15(2):249-59.  
[PUBMED](#) | [CROSSREF](#)
31. MacMahon H, Engelmann R, Behlen FM, Hoffmann KR, Ishida T, Roe C, et al. Computer-aided diagnosis of pulmonary nodules: results of a large-scale observer test. *Radiology* 1999;213(3):723-6.  
[PUBMED](#) | [CROSSREF](#)
32. Quekel LG, Kessels AG, Goei R, van Engelshoven JM. Detection of lung cancer on the chest radiograph: a study on observer performance. *Eur J Radiol* 2001;39(2):111-6.  
[PUBMED](#) | [CROSSREF](#)
33. Monnier-Cholley L, Carrat F, Cholley BP, Tubiana JM, Arrivé L. Detection of lung cancer on radiographs: receiver operating characteristic analyses of radiologists', pulmonologists', and anesthesiologists' performance. *Radiology* 2004;233(3):799-805.  
[PUBMED](#) | [CROSSREF](#)
34. Del Ciello A, Franchi P, Contegiacomo A, Cicchetti G, Bonomo L, Larici AR. Missed lung cancer: when, where, and why? *Diagn Interv Radiol* 2017;23(2):118-26.  
[PUBMED](#) | [CROSSREF](#)