Article 25fa pilot End User Agreement

This publication is distributed under the terms of Article 25fa of the Dutch Copyright Act (Auteurswet) with explicit consent by the author. Dutch law entitles the maker of a short scientific work funded either wholly or partially by Dutch public funds to make that work publicly available for no consideration following a reasonable period of time after the work was first published, provided that clear reference is made to the source of the first publication of the work.

This publication is distributed under The Association of Universities in the Netherlands (VSNU) ‘Article 25fa implementation’ pilot project. In this pilot research outputs of researchers employed by Dutch Universities that comply with the legal requirements of Article 25fa of the Dutch Copyright Act are distributed online and free of cost or other barriers in institutional repositories. Research outputs are distributed six months after their first online publication in the original published version and with proper attribution to the source of the original publication.

You are permitted to download and use the publication for personal purposes. Please note that you are not allowed to share this article on other platforms, but can link to it. All rights remain with the author(s) and/or copyrights owner(s) of this work. Any use of the publication or parts of it other than authorised under this licence or copyright law is prohibited. Neither Radboud University nor the authors of this publication are liable for any damage resulting from your (re)use of this publication.

If you believe that digital publication of certain material infringes any of your rights or (privacy) interests, please let the Library know, stating your reasons. In case of a legitimate complaint, the Library will make the material inaccessible and/or remove it from the website. Please contact the Library through email: copyright@ubn.ru.nl, or send a letter to:

University Library
Radboud University
Copyright Information Point
PO Box 9100
6500 HA Nijmegen

You will be contacted as soon as possible.
The Effect of Supplementary Bone-Suppressed Chest Radiographs on the Assessment of a Variety of Common Pulmonary Abnormalities

Results of an Observer Study

Steven Schalekamp, MD,* Nico Karssmeijer, PhD,* Ariane M. Cats, MD,† Bartjan De Hoop, MD, PhD,* Bram H.J. Geurts, MD,* Olette Berger-Hartog, MD,* Bram van Ginneken, PhD,* and Cornelia M. Schaefer-Prokop, MD, PhD*†

Purpose: The aim of the study was to investigate the effect of bone-suppressed chest radiographs on the detection of common chest abnormalities.

Materials and Methods: A total of 261 posteroanterior and lateral chest radiographs were collected from 2 hospitals. Radiographs could contain single or multiple focal opacities <3 cm (n = 66), single or multiple focal opacities >3 cm (n = 33), diffuse lung disease (n = 49), signs of cardiogenic congestion (n = 26), or no abnormalities (n = 110). Twenty-one cases contained >1 type of disease. All abnormalities were confirmed by a computed tomographic scan obtained within 4 weeks of the radiograph. Bone-suppressed images (BSIs) were generated from every posteroanterior radiograph (ClearRead BSI 3.2). All cases were read by 6 radiologists without BSI, followed by an evaluation of the same case with BSI. Presence or absence of each disease category and confidence (0-100) of the observers were documented for each interpretation. Differences in the number of correct detections without and with BSI were analyzed using the Wilcoxon signed-rank test.

Results: On average, 6 more cases with focal lesions were correctly identified with BSI ($P = 0.03$), and 1 additional case with diffuse abnormalities was found with BSI ($P = 0.32$). None of the observers demonstrated a decrease in the number of correctly detected cases with diffuse abnormalities or cardiogenic congestion with BSI. False positives in normal cases with availability of BSI detected cases with diffuse abnormalities (n = 110). Twenty-one cases contained >1 type of disease. All abnormalities were confirmed by a computed tomographic scan obtained within 4 weeks of the radiograph. Bone-suppressed images (BSIs) were generated from every posteroanterior radiograph (ClearRead BSI 3.2). All cases were read by 6 radiologists without BSI, followed by an evaluation of the same case with BSI. Presence or absence of each disease category and confidence (0-100) of the observers were documented for each interpretation. Differences in the number of correct detections without and with BSI were analyzed using the Wilcoxon signed-rank test.

Conclusions: BSI does not negatively affect the interpretation of diffuse lung disease, while improving visualization of focal lesions on chest radiographs. BSI leads to overcalling of focal abnormalities in normal radiographs.

Key Words: bone-suppressed images, chest radiography, common chest abnormalities, focal chest abnormalities, observer study

(J Thorac Imaging 2016;31:119–125)
Therefore, we undertook an observer study to investigate the effect of BSIs on the detection of the most common types of pulmonary abnormalities on chest radiographs. We included cases with focal opacities of variable size, cases with diffuse lung disease or signs of cardiogenic congestion, and normal cases. To be able to focus on detection performance rather than on interpretation, we defined disease categories. For each case, radiologists scored the presence of 1 or multiple of the 4 disease categories without and subsequently with BSIs. The objective of the study was to measure the effect of reading with the supplementary BSI on the detection of focal abnormalities in a setting with a variety of common pulmonary abnormalities, and the potential to improve the detection of other pulmonary abnormalities.

MATERIALS AND METHODS

Data

Patient informed consent was waived by our institution’s ethics committee for this retrospective analysis of existing radiographic data. Study images were collected from 2 institutions (Radboud University Medical Center Nijmegen and Meander Medical Center Amersfoort, the Netherlands). Chest radiographs in both PA and lateral projections needed to be available for inclusion in this study. All abnormalities on the chest radiographs, as well as the absence of abnormalities on normal chest radiographs, were confirmed by a coherent computed tomographic (CT) scan obtained within 4 weeks of the chest radiograph. CT findings served as reference for the presence or absence of disease, upon which the visibility of corresponding findings on radiographs was confirmed by the researcher and an experienced chest radiologist in consensus. Radiographs could contain single or multiple focal opacities smaller or larger than 3 cm, ILM, or signs of cardiogenic congestion. Radiographs could contain only 1 type of disease or multiple types of diseases. We aimed to include cases with a wide range of subtleties and to have a balanced distribution between cases with focal and diffuse disease. Radiographs could show signs of chronic obstructive pulmonary disease (COPD); however, readers were instructed not to weight COPD as a separate disease entity. All radiographs were reviewed and categorized into a disease category by the clinical researcher and an expert radiologist (>15 y of experience). Classification into disease categories (focal opacities <3 cm, focal opacities >3 cm, diffuse disease, cardiogenic congestion) was based on the abnormality distribution on the chest radiograph and the phenotype on CT. All images were blinded with respect to patient demographics.

Image Acquisition

Chest radiographs were acquired with storage phosphor radiography (CR; Agfa, Mortsel, Belgium), flatpanel direct radiography (Siemens, Erlangen, Germany), and slot scanning charge coupled device technology (Delf Diagnostic Imaging, Veenendal, the Netherlands). Image processing was applied as recommended by the manufacturers and in use for clinical routine in the institutions. For all patients PA and lateral projections were available.

Bone Suppression Software

BSIs were generated by ClearRead Bone Suppression 3.2 (Riverain Technologies, Miamisburg, OH). This most recent software version suppresses all bony structures in the frontal chest radiograph and is designed to produce an image that has the same characteristics as the original image with respect to gradation, detail contrast, and size. The software product has US Food and Drug Administration approval.

Observer Study

Six observers (4 radiologists and 2 residents) from 2 hospitals participated in the study. A sequential reading design was used for reading of the images, meaning that first the original PA and lateral radiographs were evaluated, immediately followed by a second evaluation of the same case with additional availability of the PA BSI. Per case, without and with BSI, the observers were asked to fill in a score form. The observers reported the presence or absence of the previously mentioned disease categories, without mentioning underlying pathology or pathophysiology. The 4 disease categories were defined as follows: focal opacities smaller than 3 cm, focal opacities larger than 3 cm, diffuse diseases (eg, ILD/airways disease; present in all lung fields), and signs of cardiogenic congestion (including various degrees of pulmonary edema, and could include minor pleural fluid and Kerley B-lines). Observers only had to score whether one or more of these 4 disease categories were present; they did not have to locate the abnormality. Finally, per case observers were asked to quantify their confidence in their evaluation on a continuous scale between 0 and 100 without and with BSI. Readers were not informed about the distribution of normal and abnormal cases or about the frequency of disease categories in the study group.

For the evaluation of the chest radiographs we used an in-house-developed workstation (Blinded). This workstation allows for digital documentation of score forms. The observers evaluated all cases in different randomized orders having the PA and the lateral radiograph available. The PA BSIs were stacked at the same position as the original PA radiograph on the same monitor. Observers could toggle between the PA radiograph and the PA BSI to review corresponding areas in the images. Commonly used processing tools such as adjustment of window and level, and grayscale inversion, were available in the workstation.

We used a 30-inch LCD monitor (Flexscan SX3031W; Eizo, Ishikawa, Japan; native screen resolution of 2560 by 1600) for visualization of the images. Readings took place under dimmed light conditions.

Observers were told that none, 1, or >1 disease category could be present per case. Readers quantified their confidence into detection of diseases per case and not per disease category. Detection scores and confidence scales were digitally documented first without BSI and subsequently with the availability of BSI. Any given score could not be changed retrospectively. Observers were instructed to disregard focal linear opacities (corresponding to plate-like atelectasis or scars) and calcified granulomas and not to score COPD as diffuse abnormality. A training set of 20 cases to become familiar with the type of abnormalities present in the study group and the effects of BSI was provided to the observers. During this training the observer received instant feedback from the clinical researcher. Two of the 6 observers had previous experience with the bone suppression software from previous studies; none of the observers had used BSI in clinical practice.
Data Analysis

The number of correct classifications per disease category and the number of false positives in normal cases were calculated per observer. Analysis focused on change in reported disease categories on the score form and confidence of the interpretation with and without the use of BSI. Number and type of cases in which the opinion of the observer changed on the basis of the BSI were documented. Increase and decrease in confidence in cases without and with change in disease categorization were measured per case and for all cases per observer. Significance of differences per reader with respect to confidence scores was calculated using a paired t test. The number of correctly classified diseased categories without and with BSI was compared using the Wilcoxon signed-rank test. The significance level was set at $P < 0.05$. Statistical analysis was performed with IBM SPSS Statistics (version 20.0).

RESULTS

Patient Characteristics

In total 261 radiographs were collected: 153 radiographs were from male patients and 108 radiographs from female patients. The average age of the total group was 55.5 ($\pm$ 16.8) years. The average age of patients with diseased radiographs was 58.0 ($\pm$ 14.8) years, compared with an average age of patients with normal chest radiographs ($P = 0.05$, $t$ test). Men and women were evenly distributed over the diseased and the normal group (male:female; diseased 88:63, normal 65:45). The distribution

![Figure 1](image1.png)  
**FIGURE 1.** Example of a case with multiple small lesions; a 51-year-old man with multiple small focal lesions based on metastasis, correctly observed by all 6 readers. Original radiograph (A); BSI (B).

![Figure 2](image2.png)  
**FIGURE 2.** Example of a case with a large focal lesion; a 76-year-old man with a large focal lesion in the right lung that was reported by 4 of the 6 readers. Original radiograph (A); BSI (B).
for each disease category was as follows: 66 cases with single or multiple small focal opacities (<3 cm), 33 cases with single or multiple large focal opacities (>3 cm), 49 cases with diffusely distributed lung pathology, 26 cases with signs of cardiogenic congestion, and 110 normal cases (Figs. 1–4). Twenty-one cases showed >1 disease category.

Impact on Focal Disease

Without BSI, 49 (range, 29 to 59) of the 99 focal lesions were reported on average by the observers. This number increased to 55 when BSIs were used ($P = 0.03$). Especially smaller lesions (<3 cm) were more easily seen with BSI (on average 5 additional detections, $P = 0.03$) compared with large focal lesions (on average 0.5 additional detections, $P = 0.08$). All observers found more small focal lesions with BSI, and none of the observers demonstrated a decrease in the number of focal lesions with BSI (Table 1).

Impact on Diffuse Disease

On average, 42 (range, 29 to 55) of the 65 cases were correctly identified with diffuse abnormalities (including cardiogenic congestion) (Fig. 3).
cases with cardiogenic congestion) on the basis of the original radiograph. On average, 1 extra case with diffuse abnormalities was found with BSI. There was no significant difference in the number of reported cases without or with BSI for diffuse disease \((P=0.32)\) or cardiogenic congestion \((P=0.16)\). None of the observers demonstrated a decrease in the number of correctly detected cases with diffuse abnormalities or cases with signs of cardiogenic congestion (Table 1).

**Evaluation of Normal Cases**

Of the 110 normal cases in this study, each reader reported on average 27 (25%) as abnormal, based solely on the original radiograph. The most frequent falsely reported abnormalities in the normal group were small focal lesions \((n=11)\) and diffuse lung disease \((n=11)\).

With BSI, on average 34 cases were reported as abnormal, 7 more than without BSI \((P=0.04)\). With BSI, especially the number of false-positive reports on small focal lesions increased \((n=18)\), whereas the number of false-positive reports on diffuse lung disease \((n=10)\) decreased. Only 1 observer did not report more false-positive findings in normal cases with the use of BSI. An overview of changes in normal cases with BSI is shown in Table 2.

**Changes in Confidence**

One of the observers showed a significant decrease in the confidence level with BSI (from 92 to 90, \(P=0.01\)). Two observers showed a significant increase in their confidence level with BSI (from 79 to 84, \(P<0.001\) and from 74 to 76, \(P<0.001\)). The remaining 3 observers showed no significant difference in confidence level with and without BSI.

Change in confidence was dependent on a change in disease categorization. We found an increase in confidence in those cases in which the observers did not change the reported disease category. A decrease in confidence, however, was observed in cases in which observers reported a different disease category with the BSIs, regardless of whether the change was correct with respect to the disease reference (Table 3).

**DISCUSSION**

In this article we showed that BSIs improve detection of focal lesions without negatively affecting the detection of diffuse chest abnormalities and cardiogenic congestion. None of the observers demonstrated a decrease in the number of correctly detected diffuse chest abnormalities and cardiogenic congestion with the use of BSIs. At the same time, we did find a significant increase in the detection of focal opacities, especially for the smaller lesions. This beneficial effect of BSI, however, was counterbalanced by the fact that BSI invoked 7 false-positive results in normal patients, all caused by false observation of focal lesions (Fig. 5). Most interestingly, BSI did not result in an overcall of diffuse lung disease in normal patients.

For the use of BSI in clinical practice, it is important to know for which type of pathology the software has beneficial effects and for which type of disease BSI has detrimental effects. Previous studies discussed the detection of focal lesions,1-3 but thus far none has investigated the effect of BSI on a set of cases with mixed abnormalities, including diffuse lung disease. It is known that only a small

### TABLE 1. Reported Disease Categories

<table>
<thead>
<tr>
<th>Disease Category</th>
<th>n</th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
<th>6</th>
<th>Average</th>
</tr>
</thead>
<tbody>
<tr>
<td>Focal lesions &lt; 3 cm</td>
<td>66</td>
<td>35</td>
<td>34</td>
<td>46</td>
<td>42</td>
<td>37</td>
<td>26</td>
<td>36.7</td>
</tr>
<tr>
<td>Focal lesions &gt; 3 cm</td>
<td>33</td>
<td>40</td>
<td>45</td>
<td>49</td>
<td>50</td>
<td>39</td>
<td>28</td>
<td>41.8*</td>
</tr>
<tr>
<td>Diffuse</td>
<td>49</td>
<td>18</td>
<td>16</td>
<td>13</td>
<td>6</td>
<td>17</td>
<td>3</td>
<td>12.7</td>
</tr>
<tr>
<td>Cardiogenic congestion</td>
<td>26</td>
<td>19</td>
<td>17</td>
<td>14</td>
<td>6</td>
<td>17</td>
<td>3</td>
<td>12.7</td>
</tr>
<tr>
<td>Normal</td>
<td>110</td>
<td>75</td>
<td>64</td>
<td>92</td>
<td>63</td>
<td>61</td>
<td>102</td>
<td>76.2*</td>
</tr>
</tbody>
</table>

Number of correctly reported disease categories without and with BSI, for each observer.

*Significant difference between the number of correctly classified cases per disease category without and with BSI.

### TABLE 2. Reported Abnormalities in Normal Patients

<table>
<thead>
<tr>
<th></th>
<th>Obs 1</th>
<th>Obs 2</th>
<th>Obs 3</th>
<th>Obs 4</th>
<th>Obs 5</th>
<th>Obs 6</th>
<th>Average</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>No BSI</td>
<td>With BSI</td>
<td>No BSI</td>
<td>With BSI</td>
<td>No BSI</td>
<td>With BSI</td>
<td>No BSI</td>
</tr>
<tr>
<td>Focal &lt; 3 cm</td>
<td>15</td>
<td>12</td>
<td>13</td>
<td>30</td>
<td>10</td>
<td>13</td>
<td>16</td>
</tr>
<tr>
<td>Focal &gt; 3 cm</td>
<td>1</td>
<td>2</td>
<td>4</td>
<td>7</td>
<td>3</td>
<td>4</td>
<td>2</td>
</tr>
<tr>
<td>Diffuse</td>
<td>15</td>
<td>13</td>
<td>10</td>
<td>8</td>
<td>1</td>
<td>1</td>
<td>20</td>
</tr>
<tr>
<td>Cardiogenic congestion</td>
<td>1</td>
<td>0</td>
<td>1</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
</tbody>
</table>
portion of the chest radiographs in clinical practice show incidental solitary lung nodules, or actionable lung nodules in screening studies. Thus, as in a clinical setting a large number of chest radiographs will not contain lung nodules, it is important to estimate the effect of BSI on the evaluation of other common chest abnormalities as well as on the evaluation of normal cases.

We were concerned whether increased visualization of the lung architecture would result in an overcall of diffuse lung disease. The results of our observer study revealed that BSI had no detrimental effect on the evaluation of diffuse parenchymal changes: we found considerable variability in the number of reported cases with ILD, but these numbers were not affected by BSI.

We observed a substantial overcall of small focal lesions in normal radiographs (11 on average). With BSIs this number further increased to 18. Overcalling of focal lesions in normal cases will cause unnecessary follow-up. One reason for the increased number of reported focal lesions in normal cases could be suboptimal processing resulting in artifacts that mimic focal lesions. Another reason could be the increased visibility of vascular crossings. BSIs are not developed to replace the chest radiographs and therefore should always be used as adjunct to the original radiograph; abnormalities seen on BSI have to be verified on the original radiograph to avoid false-positive calls. It is very likely that more experience with the software will at least partially overcome this problem: None of the observers had used BSI beyond study conditions. Furthermore, it has to be noted that the case set was enriched with radiographs containing focal lesions, and therefore readers may have reported focal abnormalities more aggressively than they would do in clinical practice. Finally, the sequential reading design attracts readers to point out more lesions in the second evaluation step with BSI. Although no significant bias for the sequential reading methodology using receiver operating characteristic statistics is established, these absolute numbers of reported cases with small focal abnormalities may not reflect the clinical situation.

Categorization of different types of lung diseases is somewhat artificial, and radiologists are not used to categorizing the multiple findings especially seen in radiographs of elderly patients (average age 56 y) into distinct, non-overlapping categories. In addition, the reader behavior is influenced by study conditions to an unknown extent. The

<table>
<thead>
<tr>
<th>Observer</th>
<th>No. Cases Without Change in Disease Category (Change in Confidence)</th>
<th>No. Cases With Change in Disease Category (Change in Confidence)</th>
<th>From False to Correct Disease Category</th>
<th>From Correct to False Disease Category</th>
<th>From False to False Disease Category</th>
</tr>
</thead>
<tbody>
<tr>
<td>Observer 1</td>
<td>218 (+ 0.9)</td>
<td>43 (− 6.6)</td>
<td>14</td>
<td>8</td>
<td>21</td>
</tr>
<tr>
<td>Observer 2</td>
<td>211 (+ 4.9)</td>
<td>50 (+ 3.3)</td>
<td>14</td>
<td>18</td>
<td>18</td>
</tr>
<tr>
<td>Observer 3</td>
<td>237 (+ 0.5)</td>
<td>24 (− 29)</td>
<td>7</td>
<td>7</td>
<td>10</td>
</tr>
<tr>
<td>Observer 4</td>
<td>218 (+ 0.4)</td>
<td>43 (− 7.8)</td>
<td>9</td>
<td>9</td>
<td>25</td>
</tr>
<tr>
<td>Observer 5</td>
<td>244 (+ 2.5)</td>
<td>17 (+ 1.2)</td>
<td>3</td>
<td>6</td>
<td>8</td>
</tr>
<tr>
<td>Observer 6</td>
<td>255 (+ 0.1)</td>
<td>6 (− 13)</td>
<td>2</td>
<td>0</td>
<td>4</td>
</tr>
</tbody>
</table>

Changes in disease category reported on the score forms by the observers. The number in the parenthesis represents the average change in confidence level (on a scale from 0 to 100).

FIGURE 5. Example of a normal radiograph of a 62-year-old woman that invoked false-positive calls of focal lesions with BSI. Original radiograph (A); BSI (B).
latter probably also contributed to the overcalling of focal lesions in normal radiographs. There was considerable difference between readers with respect to the use of disease categories. Therefore, we focused our analysis on individual changes introduced by BSI and not on absolute performance measures.

For this study, we tried to select a set of clinically representative cases. Although all cases were derived from clinical archives, the case set was not consecutive but enriched with abnormalities; also other abnormalities like bone lesions and pleural lesions were not included in this study. Especially the first aspect has to be considered when transferring the results into a clinical situation. To capture the full range of pathology to be encountered in clinical radiographs at realistic incidences, a very large study group would have been necessary to allow for a meaningful data analysis. We therefore decided for a case selection and disease categorization.

In summary, in this study we showed that BSIs do not negatively influence the interpretation of chest radiographs by radiologists regarding the evaluation of diffuse lung disease and signs of cardiogenic congestion. Radiologists are aided by BSI in cases with focal abnormalities; however, the availability of BSIs causes overcalling of focal opacities in normal radiographs, especially if experience with BSI is still limited.

REFERENCES