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.
Bone suppressed images improve radiologists' detection performance for pulmonary nodules in chest radiographs

Steven Schalekamp a,*, Bram van Ginneken a,1, Louis Meiss b,2, Liesbeth Peters-Bax a,1, Lorentz G.B.A. Quekel b,2, Miranda M. Snoeren a,1, Audrey M. Tiehuis b,2, Rianne Wittenberg b,2, Nico Karssmeijer a,1, Cornelia M. Schaefer-Prokop a,b,1,2

a Radboud University Nijmegen Medical Centre, Geert Grooteplein 10, 6525 GA Nijmegen, The Netherlands
b Meander Medical Centre, Utrechtseweg 200, 8800 BM Amersfoort, The Netherlands

A R T I C L E   I N F O

Article history:
Received 3 July 2013
Received in revised form 16 September 2013
Accepted 17 September 2013

Keywords:
Lung/radiography
Diagnosis, computer-assisted Radiography, thoracic/methods
Solitary pulmonary nodule/diagnosis
Solitary pulmonary nodule/radiography
Area under curve

A B S T R A C T

Objectives: To assess the effect of bone suppression imaging on observer performance in detecting lung nodules in chest radiographs.

Materials and methods: Posteroanterior (PA) and lateral digital chest radiographs of 111 (average age 65) patients with a CT proven solitary nodule (median diameter 15 mm), and 189 (average age 63) controls were read by 5 radiologists and 3 residents. Conspicuity of nodules on the radiographs was classified in obvious (n=32), moderate (n=32), subtle (n=29) and very subtle (n=18). Observers read the PA and lateral chest radiographs without and with an additional PA bone suppressed image (BSI) (ClearRead Bone Suppression 2.4, Riverain Technologies, Ohio) within one reading session. Multi reader multi case (MRMC) receiver operating characteristics (ROC) were used for statistical analysis.

Results: ROC analysis showed improved detection with use of BSI compared to chest radiographs alone (AUC = 0.883 versus 0.855; p = 0.004). Performance also increased at high specificities exceeding 80% (pAUC= 0.136 versus 0.124; p = 0.0007). Operating at a specificity of 90%, sensitivity increased with BSI from 66% to 71% (p = 0.0004). Increase of detection performance was highest for nodules with moderate and subtle conspicuity (p = 0.02; p = 0.03).

Conclusion: Bone suppressed images improve radiologists' detection performance for pulmonary nodules, especially for those of moderate and subtle conspicuity.

© 2013 Elsevier Ireland Ltd. All rights reserved.

1. Introduction

Chest radiography (CXR) represents the most frequently performed radiological examination and – despite its inherent limitation caused by overprojection – is up to now still the first line method for the diagnosis of lung nodules. Although most CXR are not made for the purpose of detecting nodules, nodules can be present and should be reported given the fact that solitary lung nodules, in the size range detectable by CXR, harbor a high risk to represent early stage lung cancer [1,2].

While early detection of lung cancer is beneficial for survival [3], it is documented that a substantial number of bronchogenic tumors are missed on the initial radiograph even though they are visible in retrospect. The actual frequency of lung cancers missed at chest radiography differs widely ranging from 19% to 26% according to the literature [1,2,4–6]. Those numbers are based on outdated radiographic technique suffering from a lower image quality, but also recent publications using modern digital radiographic technique reported comparable results, and thus confirm that the problem of misdiagnosis of early lung cancer on radiographs is still present today [7,8].

Obstruction of lung cancer by overlying bone structures represents a major contributor to misdiagnosis: Monnier et al. described an at least partial obstruction by bones in 60% of 30 missed lung cancers [5]; Shah reported that 22% of 40 missed lesions were obscured by the clavicle and 95% of the missed lesions were obscured by bones [6]. Improved focal lesion detection by subtracting the bone structures from the CXR has been realized by dual energy technique. Despite several publications that showed an

0720-048X/$ – see front matter © 2013 Elsevier Ireland Ltd. All rights reserved.
http://dx.doi.org/10.1016/j.ejrad.2013.09.016

* Corresponding author at: Department of Radiology, Route 767, Radboud University Nijmegen Medical Centre, Geert Grooteplein 10, Internal Postal Code 766, 6525 GA Nijmegen, The Netherlands. Tel.: +31 24 36 16651.
E-mail addresses: steven.schalekamp@gmail.com (S. Schalekamp), b.vanginneken@rad.umcn.nl (B. van Ginneken), L.Meiss@meandermc.nl (L. Meiss), L.Petersbax@rad.umcn.nl (L. Peters-Bax), Lgbaquekel@meandermc.nl (L.G.B.A. Quekel), m.snoeren@rad.umcn.nl (M.M. Snoeren), am.tiehuis@meandermc.nl (A.M. Tiehuis), rianne.wittenberg@hotmail.com (R. Wittenberg), n.karssmeijer@rad.umcn.nl (N. Karssmeijer), cornelia.schaeferprokop@gmail.com (C.M. Schaefer-Prokop).

1 Tel.: +31 24 16651.
2 Tel.: +31 33 850 5050.
increased lung nodule detection performance [9–13] the technique never found its way into broad clinical application. This is likely due to the special hardware requirements and an image quality that is variably limited by subtraction artifacts and image noise.

A new software product has been developed to suppress ribs and clavicles in the original image without requiring special hardware or increase of patient dose caused by double exposures. A recent study provided evidence for improved detection of lung cancer using this bone suppression software [14]. Sensitivity increased from 49.5% to 66.3% when evaluation of the CXR was aided by bone suppressed images (BSI). However, the increase in sensitivity was accompanied by a significant loss of specificity from 96.1% to 91.8% [14]. Moreover, the study had three limitations. First, only PA radiographs were assessed. Second, the majority of the study images were secondary digitized conventional radiographs, which might have lead to an underestimation of the potential of chest radiography alone. Third, at clinically relevant operating points, namely at low false positive rates, LROC analysis did not show a convincing improvement of detection performance using BSI.

Addressing these shortcomings, we therefore decided to undertake another study to assess the effect of a new version of BSI on the detection of nodules in digital chest radiographs, using PA and lateral projections.

2. Materials and methods

To assess the effect of bone suppressed images on the detection of pulmonary nodules, we conducted an observer study involving five radiologists and three residents from two institutions. Selection of study images and study setup were reviewed and approved by the institutional review board.

2.1. Image selection

To construct a study database, images were retrospectively retrieved from image archives of four hospitals (three academic and one non-academic hospital). All images had been obtained for clinical purposes. Patients with a solitary pulmonary nodule between 5 and 35 mm in diameter that underwent both posteroanterior (PA) and lateral chest radiograph as well as a chest computed tomography (CT) scans within 3 months time interval were included. All patients (diseased and controls) were older than 40 years; none of them showed other diseases than signs of chronic obstructive lung disease. It should be noted that cases were not selected with respect to nodule location (e.g. behind a rib shadow) or image quality to favor rib suppression technique. An expert radiologist (CSP) and the researcher (SS), who both did not participate in the observer study, classified in consensus the nodules in four categories of conspicuity, ranging from obvious (cat. 1) to moderate (cat. 2), subtle (cat. 3) and very subtle (cat. 4). In the study group all four categories of conspicuity were adequately represented to allow for statistically meaningful analysis of the subgroups of nodules. Overlap of bone structures was visually quantified by the researcher (SS) into minimal (<50% of lesion area), and partial to complete (50–100%).

Location of the nodule in the PA radiograph was annotated using thick coronal multiplanar reconstruction of the CT data set for correlation. Nodules that were not visible on the PA radiograph as determined in consensus by the expert radiologist and the researcher were not included into the study group. It is noteworthy that at the time of image selection no BSI was available. Randomly selected chest radiographs of patients without nodules, as proven by CT examinations obtained within 6 months of the CXR, served as controls.

Patient age, gender, smoking history and lesion histology were derived from the clinical records. Nodule volume and diameter were automatically measured on CT [15]. The study group consisted of 300 subjects: 111 patients with a solitary nodule and 189 control subjects. An additional 40 subjects, including 22 patients with a solitary nodule, were used as training cases.

2.2. Image acquisition

All chest radiographs were obtained with digital technique using storage phosphor plates (CR, Agfa, Mortsel, Belgium), Selenium drum (Thoravision, Philips, Hamburg, Germany) and flat panel detector DR systems (Siemens, Erlangen, Germany; Hologic Inc., Bedford, MA, USA). Radiographs were obtained using 120 or 125 kVp, respectively, with automatic exposure control. Image processing was applied as recommended by the manufacturers, and in use for clinical routine in the various institutions. For all patients posteroanterior and lateral projections were available.

2.3. Software

Bone suppressed images were generated by ClearRead Bone Suppression 2.4, Riverain Technologies (Miamisburg, OH). This visualization software uses advanced image processing to construct a bone suppressed PA chest image. The software 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 Food and Drug Administration (FDA) approval. Since bone suppressed images were generated in a post-processing step, no special hardware or double exposures like in dual energy radiography were needed.

2.4. Reading methodology

Eight readers, including five radiologists (experience 3–17 years) and three residents (one second year resident, and two third year residents), evaluated the 300 study cases in different randomized orders. Before the actual reading sessions, all readers evaluated a training set of 40 cases with and without BSI. During this training session, the researcher gave instantaneous feedback to the observers. None of the observers had previous experience in reading chest radiographs with BSI or dual energy chest radiography.

Readings were carried out using a 30-in. (Flexscan SW3031W; Eizo, Ishikawa, Japan) DICOM-calibrated LCD monitor in a darkened room, mimicking clinical reading conditions. The screen was large enough to review both PA and lateral radiograph side-by-side. Processing tools were available, including zoom in/out, adjustment of window and level and gray scale inversion.

Observers reviewed the PA and lateral chest radiographs first without and then immediately with availability of the BSI within one reading session. They were asked to localize and score areas suspicious for representing a nodule using a continuous scale between 0 and 100 to indicate their degree of suspicion (0=not suspicious, 100=definitely suspicious). Both, lesion location and degree of suspicion were recorded digitally. The readers first determined a score for the assessment of the CXR alone, immediately followed by a second score for the assessment of CXR and BSI. Readers were allowed to score multiple regions per image. Readers were asked to purely focus on the presence of a nodule and not to take into consideration the potential underlying diagnosis (malignancy versus benign disease) or the potential consequences (CXR follow up or CT). The BSI was displayed on the same monitor as the original PA radiographs. The readers could toggle between the original CXR and the BSI using a key on the keyboard.

Observers were informed that radiographs could contain none or a solitary nodule, and that nodules were of varying conspicuity.
### Table 1
Nodule characteristics. Categories of conspicuity, nodule size, and superprojection by bones.

<table>
<thead>
<tr>
<th>Conspicuity</th>
<th>Number</th>
<th>Average diameter in mm (SD)</th>
<th>Superprojection by bones</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>1 (obvious)</td>
<td>2 (moderate)</td>
<td>3 (subtle)</td>
</tr>
<tr>
<td>Number</td>
<td>32</td>
<td>32</td>
<td>29</td>
</tr>
<tr>
<td>Average diameter in mm (SD)</td>
<td>17.0 (5.0)</td>
<td>16.9 (6.4)</td>
<td>15.1 (5.3)</td>
</tr>
<tr>
<td>Superprojection by bones</td>
<td>&lt;50%</td>
<td>10</td>
<td>13</td>
</tr>
<tr>
<td></td>
<td>&gt;50%</td>
<td>22</td>
<td>19</td>
</tr>
</tbody>
</table>

SD = standard deviation.

They did not know the percentage of diseased and control subjects and the distribution of conspicuity.

### 2.5. Statistics

For statistical analysis, multi reader multi case (MRMC) receiver operating characteristic (ROC) analysis was performed, assigning to each case the score of the most suspicious finding. Differences in observer performance were tested using the Dorffman, Berbaum and Metz method (DBM MRMC package) which includes case, reader and treatment variance into analysis, and allows for analysis of partial area under the curve (pAUC) [16–18]. The proper binormal model was used to create ROC curves. Areas under the ROC curves (AUC) without and with use of bone suppressed images were compared for all lesions and for subsets of nodules according to their conspicuity and the degree of bone overlap. Besides complete AUC also pAUC at an interval between 0 and 0.2, corresponding to a specificity range of 80–100%, were computed. From these ROC curves, individual readers’ sensitivities were calculated at a specificity of 90%.

Besides ROC analysis we also performed location based analysis. We calculated observers’ overall sensitivity and specificity using the findings with a degree of suspicion exceeding a score of 50 and taking correct lesion localization into account. By taking location into account readers were not rewarded for marking false positive locations in abnormal images. The suspiciousness score of the above 50 was arbitrarily chosen, with the underlying idea that the nodules of which, the readers thought they were above average suspicion would require follow up. A paired t-test was used to calculate significance between sensitivities and specificities without and with BSI.

Significance of difference was defined at \( p < 0.05 \). Observers’ markings were considered true positive when the marker was within ground-truth annotation of the lesion. Differences in patient characteristics were tested using an unpaired t-test and the chi-square test.

### 3. Results

#### 3.1. Patient demographics

No significant differences were found between diseased \( n = 111 \) and control \( n = 189 \) patients with respect to mean age (64.9 versus 63.4 years \( p = 0.22 \)), and sex \( m:f = 66:45 \) versus 111:78 \( p = 0.9 \), but there was a difference in smoking history (86% versus 73% \( p = 0.004 \)).

#### 3.2. Nodule characteristics

Average nodule diameter was 16 mm (7.8–34.9 mm), with a median of 15 mm. Five nodules, exceeded a diameter of 30 mm and could be classified as mass according to the Fleischner Society Guidelines [19]. 80% of the nodules were malignant, which was histologically proven in 85% and based on clinical history in 15%.

With respect to location on the PA radiograph there were 25 apical nodules located cranial of the aortic arch, 15 (peri) hilar nodules, 9 retrocardiac nodules, and 62 peripheral nodules located in the remaining lung fields. Of all lesions, 31% \( n = 35 \) were minimally and 69% were partially to completely \( n = 76 \) superprojected by bones. The degree of superprojection by bones was not statistically correlated with conspicuity (Table 1).

#### 3.3. Observer performance: ROC analysis

Area under the curve was 0.855 for CXR alone and significantly increased to 0.883 with BSI \( p = 0.004 \) (Fig. 1). Partial area under the curve for a specificity >80% was 0.124 without BSI compared to 0.136 with BSI \( p = 0.0007 \) (Table 2). Operating at a specificity of 90%, mean sensitivity of the readers increased from 65.7% to 71.4% \( p = 0.0004 \). All 8 readers increased their performance with the use of BSI compared to reading CXR alone, with a significant increase for reader 4.

#### 3.4. Observer performance: sensitivity and specificity

Considering only scores with a suspicion above 50, readers’ sensitivity increased on average from 59.8% to 67.6% \( p = 0.002 \). On average per reader this corresponded to 9 of 111 additional nodules, correctly localized with BSI. Examples of cases where BSI did improve the nodule detection performance of the readers are shown in Figs. 2 and 3. At this same threshold, readers on average called 17 cases false positive (range: 3–35) with CXR alone yielding a specificity of 91.0%. With BSI this number increased to 22 cases (range: 4–46) corresponding to a specificity of 88.4% \( p = 0.07 \).

#### 3.5. Performance for different subtlety categories and bone obscuration

Areas under curve (AUC) for the 4 categories of decreasing order of lesion conspicuity were 0.977, 0.920, 0.803 and 0.598, respectively. With BSI average AUCs increased to 0.980, 0.949, 0.846 and

<table>
<thead>
<tr>
<th>Observer</th>
<th>AUC Without BSI</th>
<th>AUC With BSI</th>
<th>pAUC Without BSI</th>
<th>pAUC With BSI</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>0.831</td>
<td>0.870</td>
<td>0.116</td>
<td>0.129 *</td>
</tr>
<tr>
<td>2</td>
<td>0.887</td>
<td>0.890</td>
<td>0.126</td>
<td>0.141 *</td>
</tr>
<tr>
<td>3</td>
<td>0.906</td>
<td>0.917</td>
<td>0.149</td>
<td>0.153</td>
</tr>
<tr>
<td>4</td>
<td>0.771</td>
<td>0.831 *</td>
<td>0.098</td>
<td>0.121</td>
</tr>
<tr>
<td>5</td>
<td>0.868</td>
<td>0.906</td>
<td>0.124</td>
<td>0.133</td>
</tr>
<tr>
<td>6</td>
<td>0.889</td>
<td>0.912</td>
<td>0.131</td>
<td>0.141</td>
</tr>
<tr>
<td>7</td>
<td>0.809</td>
<td>0.827</td>
<td>0.116</td>
<td>0.122</td>
</tr>
<tr>
<td>8</td>
<td>0.882</td>
<td>0.909</td>
<td>0.136</td>
<td>0.146 *</td>
</tr>
<tr>
<td>Average</td>
<td>0.855</td>
<td>0.883</td>
<td>0.124</td>
<td>0.136</td>
</tr>
</tbody>
</table>

\* \( p < 0.01 \).

\* \( p < 0.05 \).
Fig. 1. Average ROC curves of all observers. Black = without BSI; dotted black = with BSI.

Fig. 2. Example case 1. 63 year old male with a 31 mm non small cell carcinoma in the right upper lobe. The lesion is obscured by clavicle and ribs. On the bone suppressed image the abnormality becomes more conspicuous and is more easily identified. None of the observers scored the nodule with a score above 50, on the basis of the CXR alone. With BSI five of the eight observers correctly identified the nodule with a score above 50.

3.6. Missed nodules

The total number of missed lesions pooled over all 8 readers went down from 255 with CXR to 189 with CXR and BSI. On average readers missed 32 nodules with CXR alone, and 24 nodules with BSI. Only 46 of the 111 nodules were detected by all 8 observers, indicating that the greater proportion of nodules (59%) was not

0.647, respectively (Table 3). Differences reached significance for lesions of moderate (category 2) and subtle (category 3) conspicuity ($p = 0.02$ and $p = 0.03$).

Areas under curve (AUC) were 0.884 and 0.842, for nodules with minimal, and partial to complete superprojection by bones, respectively. With BSI, the average AUCs increased to 0.916 ($p = 0.003$) and 0.866 ($p = 0.05$), respectively (Table 3).

<table>
<thead>
<tr>
<th>Table 3</th>
<th>Performance for subcategories of lesions. Average area under the ROC curve without and with BSI, for subgroups of lesions with different conspicuity and degree of osseous superprojection.</th>
</tr>
</thead>
<tbody>
<tr>
<td>Lesion conspicuity</td>
<td>Without BSI</td>
</tr>
<tr>
<td>Obvious (cat. 1; $n = 32$)</td>
<td>0.977</td>
</tr>
<tr>
<td>Moderately subtle (cat. 2; $n = 32$)</td>
<td>0.920</td>
</tr>
<tr>
<td>Subtle (cat. 3; $n = 29$)</td>
<td>0.803</td>
</tr>
<tr>
<td>Very subtle (cat.4; $n = 18$)</td>
<td>0.598</td>
</tr>
<tr>
<td>Superprojection by bones</td>
<td></td>
</tr>
<tr>
<td>$&lt;50%$ ($n = 35$)</td>
<td>0.884</td>
</tr>
<tr>
<td>$&gt;50%$ ($n = 76$)</td>
<td>0.842</td>
</tr>
</tbody>
</table>
located by at least one of the eight readers. With BSI the number of nodules detected by all readers increased to 55. The number of nodules missed by four or more readers went down from $n = 34$ to $n = 22$ nodules with BSI.

Most missed nodules ($n = 255$) were located peripherally ($n = 129$), followed by nodules in the hilar regions ($n = 46$), lung apices ($n = 44$) and the retrocardiac areas ($n = 36$). In absolute numbers, the decrease of missed lesions with BSI was highest in peripheral lung areas ($n = 43$) versus the hilar region ($n = 9$), lung apices ($n = 10$), retrocardiac area ($n = 4$), respectively. The proportional decrease, however, was similar in all lung areas (8.7%, 7.5%, 5.0% and 5.6%, respectively) (Table 4).

### Table 4

Missed nodules. Nodules missed without and with BSI, based on anatomic location. Number = number of nodules, actually present in that anatomic location. AVG per reader = number of nodules missed in that area, averaged over 8 readers. Proportion = number of nodules missed, expressed as percentage of all nodules in that anatomic area.

<table>
<thead>
<tr>
<th>Location</th>
<th>Number</th>
<th>AVG per reader</th>
<th>Proportion</th>
<th>AVG per reader</th>
<th>Proportion</th>
</tr>
</thead>
<tbody>
<tr>
<td>Right lung</td>
<td>61</td>
<td>18</td>
<td>28.9%</td>
<td>13</td>
<td>21.7%</td>
</tr>
<tr>
<td>Left lung</td>
<td>50</td>
<td>14</td>
<td>28.5%</td>
<td>10</td>
<td>20.8%</td>
</tr>
<tr>
<td>Peripheral</td>
<td>62</td>
<td>14</td>
<td>22.8%</td>
<td>9</td>
<td>14.1%</td>
</tr>
<tr>
<td>Hilar</td>
<td>15</td>
<td>6</td>
<td>38.3%</td>
<td>5</td>
<td>30.8%</td>
</tr>
<tr>
<td>Apical</td>
<td>25</td>
<td>8</td>
<td>30.0%</td>
<td>6</td>
<td>25.0%</td>
</tr>
<tr>
<td>Retrocardic</td>
<td>9</td>
<td>5</td>
<td>50.0%</td>
<td>4</td>
<td>44.4%</td>
</tr>
</tbody>
</table>

## 4. Discussion

Multiple studies evaluating dual energy radiography were able to prove an improved detection of pulmonary nodules by removing overlying bone structures [9–13,20]. This positive experience stimulated the advent of software methods to digitally subtract superimposing bone structures. Advantages of such software solution are that it can be applied to any kind of digital chest radiography, does not require additional hardware equipment, eliminates motion artifacts created by two exposures (as done with dual-shot energy subtraction imaging), and does not cause additional radiation dose to the patient.
First studies evaluating this software tool reported encouraging results [10,14,21,22]. Three studies reported a significantly increase of detection rate of pulmonary nodules and AUCs, indicating that the increase of sensitivity outweighed the loss of specificity. The fourth reported an improved detection of focal pneumonias.

Our results confirm the previously reported positive effect of BSI: the number of correctly localized nodules significantly increased without a significant overall loss of specificity. In our study, all 8 readers uniformly took advantage of BSI irrespective of their differing experience.

In addition to what has been previously published about the effect of BSI on the detection of focal opacities [10,14,21,22], we analyzed in more detail, which type of nodules with respect to conspicuity and degree of osseous superprojection took advantage of the availability of BSI. Secondly we evaluated whether increase of sensitivity remained significant when high specificity levels beyond 80% were preserved. Third, we used the latest version of the bone suppression software. Finally we optimized the baseline performance of chest radiography by using digitally acquired images and providing lateral projections, which is common practice in the clinic.

With respect to conspicuity, all nodules took advantage of the availability of the BSI image with exception of the obvious nodules. Performance increase reached significance (p < 0.05) for the moderately subtle and subtle lesions. Though the relative detection increase was the largest for the very subtle nodules, no significant difference was seen, most likely due to the smaller number of nodules in this category. We could not find a correlation between the degree of bone overlay and benefit by BSI, neither a proof that nodules located in the lung apex would have more benefit than nodules in other anatomic locations of the lung. We therefore conclude that the overall reduction of anatomic noise by bone suppression contributes more to the improved detection of focal lesions than the subtraction of bones superimposed on individual lesions.

The average sensitivity with CXR and BSI of 67% is comparable to the sensitivity reported by Freedman et al. when using BSI (e.g., 66.3%). Our baseline performance, however, was substantially higher (60% in our study as opposed to 49.5% in Freedman’s). This might be the result of a somewhat higher image quality of the digital radiographs but is also likely due to the fact that our readers evaluated both PA and lateral images. We therefore consider the performance difference seen in our study, though smaller but statistically significant, as being within a clinically realistic range.

One could argue that sensitivities of 59.8% without and 67.6% with BSI are still suboptimal given the fact that all lesions were visible on the CXR with knowledge of the CT scan. The relatively low overall sensitivity underlines the fact that even under study conditions the detection of nodules – though all of them visible in retrospect – indeed represents a difficult visual task. A number of large studies have shown that the sensitivity of unaided CXR for the detection of small nodules is too low to be used as screening tool yielding significant mortality reduction [23–25]. For a number of reasons CXR is unlikely to ever achieve a sensitivity equal to that of CT. The primary goal of additional processing and display tools for chest radiography such as bone suppression or temporal subtraction is therefore to reduce the number of lesions that are missed, but retrospectively visible. These types of lesions are mostly of intermediate conspicuity and therefore especially prone to “inattentional blindness” of the observers. It is therefore an encouraging result that BSI indeed improved the detection of nodules of intermediate and subtle conspicuity and led to a uniformly improved reader performance. As a consequence the number of detected nodules by all readers increased from 46 to 55 lesions, and similarly the number of lesions missed by the majority of readers decreased from 34 to 22.

The increase of sensitivity with BSI has to outweigh the loss of specificity. Our ROC statistics found an increased sensitivity at all levels of specificity indicating that the trade-off for the improved detection rate was not a relevant loss of specificity. ROC analysis of partial area under the curve (pAUC) at high specificity levels between 80% and 100% also found a significant increase in performance. In that respect our results surpass the findings reported by Freedman et al. One could argue that readers improved their sensitivities by being able to mark several locations though they knew that the radiographs contained only one nodule. We did so to avoid the effects of satisfaction of search after having marked one possible lesion but to rather encourage the readers to scrutinize the image for potential lesions completely. A high number of false positive findings, however, would have inevitably lead to a performance decrease using ROC analysis. Thus by our statistical analysis we assured that readers could not take credit by simply marking multiple suspicious lesions.

Freedman et al. determined specificity in their study by the number of CT examinations asked for further diagnostic work-up of patients with a normal CXR: they observed a significant loss of specificity that was overruled by the larger increase of sensitivity. We determined sensitivity/specificity based on the degree of suspicion using a threshold of 50 on a scale from 0 to 100 and calling only correctly localized lesions true positive. We observed only a small and non-significant, loss of specificity (91% versus 88.4%, p = 0.07). Some of the false positives were provoked by incomplete suppression of the first rib (Fig. 4). With improvement of the suppression technique and/or more experience of the readers, decrease

---

Fig. 4. Example case 3. 83 year old female without a nodule. Incomplete suppression of the first rib produces a pseudolesion in the right upper quadrant in the BSI, which was called positive by two observers.
of false positive decisions can be anticipated. If we would assume that a lesion with a suspicion score above 50 would trigger a CT examination for further diagnostics, the number of unnecessary CTs induced by false positive lesions would have increased on average per reader from \(n = 17\) to \(n = 22\) which in a group of 189 control patients amounts to 9.0% and 11.6%. It is noted that these numbers were computed with a somewhat arbitrary threshold. The real effect of using BSI on the number of referrals to CT can only be determined in a prospective clinical study, since threshold levels and reader decisions are strongly influenced by parameters such as disease prevalence and reader vigilance because of specific search tasks.

Our study suffers from few limitations. First, we studied a selected patient group. Although all cases were derived from clinical work-up, the prevalence of nodules was substantially higher than in clinical practice. Also the fact that only one nodule per case present, might have affected the observers' behavior. Furthermore, since observations are done in a testing environment, it is less likely that observers missed nodules due to inattention, which may be an important factor in clinical practice. Finally, observers had no experience with BSI. Although a training set of 40 cases was provided, more training might be needed to learn about the strengths and weaknesses of BSI.

In conclusion, our study results found an improved detection of pulmonary nodules with use of BSI. The increase in sensitivity outweighed the non-significant loss of specificity. Detection increase was mostly seen for nodules of moderate to subtle conspicuity irrespective of the degree of superprojection by bones or lesion location. All observers irrespective of their experience reached the best performance with BSI, also when evaluation was confined to a high specificity range.

**Role of funding source**

Riverain Technologies provided a research grant to NK. First author was hired from this research grant. Riverain Technologies had no influence on the execution of the study or the decision to submit this paper for publication. Data and analysis were managed by the authors.

**References**


