Advances in digital chest radiography: impact on reader performance
De Boo, D.W.

Citation for published version (APA):

General rights
It is not permitted to download or to forward/distribute the text or part of it without the consent of the author(s) and/or copyright holder(s), other than for strictly personal, individual use, unless the work is under an open content license (like Creative Commons).

Disclaimer/Complaints regulations
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 Ask the Library: http://uba.uva.nl/en/contact, or a letter to: Library of the University of Amsterdam, Secretariat, Singel 425, 1012 WP Amsterdam, The Netherlands. You will be contacted as soon as possible.
Computed Radiography versus Mobile Direct Radiography for Bedside Chest Radiographs: Impact of Dose on Image Quality and Reader Agreement

Diederick W De Boo
Michael Weber
Eline E Deurloo
Geert J Streekstra
Nicole J Freling
Dave A Dongelmans
Cornelia M Schaefer-Prokop

Clinical Radiology 2011; 66:826-832
Abstract

Objective
To assess the image quality and potential for dose reduction of mobile direct detector (DR) chest radiography as compared with computed radiography (CR) for intensive care unit (ICU) chest radiographs (CXR).

Material and methods
Three groups of age-, weight- and disease-matched ICU patients (n=114 patients; 50 CXR per acquisition technique) underwent clinically indicated bedside CXR obtained with either CR (single read-out powder plates) or mobile DR (GOS-TFT detectors) at identical or 50% reduced dose (DR\textsubscript{50%}). Delineation of anatomic structures and devices used for patient monitoring, overall image quality and disease were scored by four readers. In 12 patients pairs of follow-up CR and DR images were available, and in 15 patients pairs of CR and DR\textsubscript{50%} images were available. In these pairs the overall image quality was also compared also side-by-side.

Results
Delineation of anatomy in the mediastinum was scored better with DR or DR\textsubscript{50%} than with CR. Devices used for patient monitoring were seen best with DR, with DR\textsubscript{50%} being superior to CR. In the side-by-side comparison, the overall image quality of DR and DR\textsubscript{50%} was rated better than CR in 96% (46/48) and 87% (52/60), respectively. Inter-observer agreement for the assessment of pathology was fair for CR and DR\textsubscript{50%} (κ = 0.33 and κ=0.39, respectively) and moderate for DR (κ = 0.48).

Conclusion
Mobile DR units offer better image quality than CR for bedside chest radiography and allow for 50% dose reduction. Inter-observer agreement increases with image quality and is superior with DR while DR\textsubscript{50%} and CR are comparable.
Introduction

Digital radiography using storage phosphor plates (computed radiography, CR) is widely available and generally accepted for both, upright and bedside chest radiography. More recently flat-panel X-ray detectors (direct radiography, DR) using amorphous Silicon thin-film-transistor with different types of converter arrays have become increasingly used for upright chest radiography. Direct radiography offers improved dose quantum efficiency compared to computed radiography. This higher dose quantum efficiency can either be used to reduce patients’ dose while maintaining image quality or to increase signal-to-noise ratio and image quality. Both features have proven validity for upright chest radiography: DR achieved an increased image quality compared with CR when acquired with the same dose\(^1,2\); at the same, a possible dose reduction of up to 60% was found without loss of diagnostically relevant image quality\(^1,3-6\). Mobile DR units use a different scintillator material (gadolinium oxysulphide = GOS) that has a lower quantum efficiency than cesium iodide = CsI, which is used for most upright radiography units\(^7,8\). Therefore, the dose reduction rates reported for upright radiography may not necessarily be transferable to bedside radiography. Conversely, from a diagnostic point-of-view image quality requirements are different at the bedside as compared with upright radiography. So far publications on mobile DR chest radiography are limited to pediatric applications. They reported a dose reduction of 50% possible with DR compared to CR without decrease of relevant image quality\(^9,10\). The aim of the present study was, therefore to, compare mobile DR (Gd\(_2\)O\(_2\)S) with mobile CR (storage phosphor plates) for adult bedside chest radiography with respect to image quality and the potential for dose reduction and the impact on diagnostic performance.
Material and Methods

Study group

One hundred and fifty anteroposterior bedside chest radiographs (CXR) were obtained of patients admitted to an adult mixed medical-surgical intensive care unit (ICU) of a large university hospital. All CXR were obtained due to clinical indication. Patient informed consent was, therefore, not necessary, and approval for the study had been obtained by the local ethic committee (071711598). The images were randomly obtained with one of the following three acquisition techniques: CR, DR or DR with 50% dose reduction (DR_{50\%}). Fifty CXRs were obtained per acquisition technique. For organizational reasons images included in the study were obtained only on selected days per week. Which patient would be radiographed depended on the clinician’s indication; the assignment to one of the three techniques was random. None of the patients underwent a CXR using different techniques on the same day. Twenty-seven patients, however, underwent several CXRs using two or three of the above mentioned techniques, leading to a study group of 150 bedside CXRs obtained in 114 patients.

Detector systems

Images were obtained using storage phosphor technique (Agfa, Mortsel, Belgium) or direct detector flat-panel technology (MOBILETT XP Digital, Siemens, Erlangen, Germany). The storage phosphor plates (Agfa) had a size of 35 x 43 cm, a pixel matrix of 2048 x 2500 with a pixel size of 0.17 mm. They were based on powder phosphor plates, and were read-out by a focal spot laser only from one side. Images were processed using multifrequency processing (MUSICA, Agfa), processing parameters were used as recommended by the manufacturer. The direct detector unit uses a scintillator (Gd_{2}O_{2}S) layered on top of an array with light-sensitive photodiodes with thin-film transistors. The 35 x 43 cm detector plate had a 2208 x 2688 pixel matrix with a pixel size of 0.16 mm. Standard post-processing available on the detector unit (Siemens) was used.

Superimposition of catheter material

The presence of monitoring devices was simulated by superimposing fragments of varying types of catheters (chest tube, central venous line, naso-gastric tube, Swan-Ganz catheter) over the upper abdomen of the patients. Thirteen templates were made containing different combinations of three to five catheter fragments fixed on the lower end of hardcopies of posteroanterior chest radiographs. The hardcopies
were used as carriers and were taped on the cassette before it was placed underneath the patient for acquisition of the anteroposterior CXR. Each fragment had a length ranging from 3 to 5 cm. The fragments were positioned in such a way as to be randomly imaged over the area of the upper abdomen. The area was chosen to avoid interference of the superimposed material with the clinical purpose of the radiograph, but to be able to test the detectability in a high attenuation area of the body in a standardized manner.

Image acquisition

Image acquisition parameters were manually set by three technicians. All images were obtained without an antiscatter grid at 90 kVp. The milliampere-second (mAs) depended on the patient’s weight and the visual assessment of the patients’ constitution by the technicians and varied from 1.0 to 1.4 for standard CR / DR and was divided by a factor of two for DR_{50\%} (0.5-0.7 mAs). All images were obtained using the same tube system, which automatically recorded dose-area product (DAP). From the DAP, patient entrance skin dose and effective dose were calculated (Table 1).

The acquisition technique used in each patient was determined by chance. Care, however, was taken during image inclusion that the three different groups were matched with respect to patient weight, body mass index (BMI) and total number of superimposed catheter fragments. For each patient weight, height, mAs, DAP and the template of superimposed catheter material used were documented. In 27 patients, pairs of follow-up images using two different acquisition techniques were available: there were 12 pairs of CR and DR images and 15 pairs of CR and DR_{50\%} images available for direct comparison.

Image evaluation

Images were evaluated in three reading sessions using a dedicated picture archiving and communication system (PACS) system (Agfa Impax, version 4.5) equipped with high resolution LCD monitors (Barco, MDCG 2121-CB, pixel matrix of 1.2K x 1.9K) that were regularly controlled for their DICOM conformity. Two radiology residents (1st and 4th year training) and two certified radiologist (both more than 15 years of experience) independently interpreted the images in different random order. Reading conditions with subdued ambient light were kept constant throughout all reading sessions. Readers were allowed to use processing tools, such as windowing or magnification, at their preference. The readers’ assessment referred to image quality criteria as well as to the diagnostic definition of the presence of disease, and compiled
the following tasks:

1. Assessment of the visibility of nine anatomic landmarks / lines using a three-point scale (3 = visible, sharply demarcated contours; 2 = moderately visible with partially blurred contours; and 1 = poorly visible with blurred contours). All anatomic landmarks were located within the high attenuation area of the chest radiograph or represented interfaces between low and high attenuation areas and included the trachea, the carina, retro-cardiac vessels, the upper / lower thoracic spine, both hemi-diaphragms, the esophageal-diaphragmatic recess and the descending aorta-diaphragmatic recess.

2. Assessment of the visibility of the superimposed catheter fragments using a similar three-point scale (3 = sharp contours throughout whole length; 2 = blurred contours but visible throughout whole length; and 1 = not visible throughout whole length).

3. Determination of the presence or absence of four different types of frequently encountered diseases in an ICU setting applying a five-point confidence scale ranging from 5 = definitely present; 4 = probably present; 3 = equivocal; 2 = probably not present; and 1 = definitely not present. The four diseases included: a) pulmonary consolidations, b) vascular congestion, c) pleural effusion and d) atelectasis. Criteria for the presence of these four pathologies followed generally accepted morphological criteria. Consolidations could be based on pneumonia, edema, or acute respiratory distress syndrome (ARDS), no specific interpretation was required. Generally, consolidations had blurred borders, could be patchy or confluent, and showed no signs of volume loss. There was no further subdivision with respect to the size of the consolidations. Conversely atelectasis were sharply demarcated, showed signs of volume loss, and were mostly located in the dependant parts of the lower lungs.

4. Assessment of the overall image quality using the Radlex ® image quality scoring criteria (11): 0 = non-diagnostic (unacceptable for diagnostic purposes); 1 = limited (acceptable, with some technical defects but still adequate for diagnostic purposes); 2 = diagnostic (acceptable, with no technical defects likely to impair using the images for diagnosis); 3 = exemplary (good, most adequate for diagnostic purposes).

All images were individually scored. If pairs of images from the same patient using two different techniques were available, images were first assessed individually and in a separate reading session in a side-by-side comparison using a forced-decision-model meaning that readers were requested to rank the images differently according to their image quality.
**Statistical analysis**

*Power analysis of statistics*

A power analysis was performed (nQuery Adviser, version 5.0, statistical solutions, Cork, Ireland) to compare the image quality scores of the different acquisition techniques. For an alpha error value of 0.05, a power of 0.82 was calculated.

*Match of patient groups*

Statistical significance of differences between the three patient groups with respect to age, BMI, acquisition dose and prevalence of disease was tested using Pearson’s Chi$^2$ test. There was no standard of truth for the definition of disease in the study group. To be able to assess differences between the three patient groups with respect to prevalence of disease, a disease category was assumed to be present if at least two of the four readers had rating scores of 4 or 5 (“probably” or “definitely”) for its presence.

*Visibility of catheter fragments and of anatomical landmarks*

The statistical significance of differences between the three acquisition techniques with respect to the visibility of catheter fragments and anatomical landmarks was assessed using an analysis of variance with repeated measures. The dependant variables in this test included the visibility scores with the three techniques: independent variables were acquisition techniques and readers. The analysis of variance tested the significance of influence of the acquisition technique and reader on the visibility scores and whether there were significant interactions between the independent variables. Hochberg’s post hoc test was calculated to determine the differences between the three acquisition techniques. Statistical analysis was carried out separately for the visibility of catheters, the visibility of the five landmarks located in the high attenuation areas of the chest (i.e., trachea, the carina, retro-cardiac vessels and the upper / lower thoracic spine) and of the four landmarks located at density interfaces (both hemi-diaphragms, the esophageal-diaphragmatic recess and the descending aorta-diaphragmatic recess). SPSS version 11.5 was used for statistical analysis.

*Subjective assessment of overall image quality*

The statistical significance of differences between the three acquisition techniques with respect to overall image quality ratings was assessed using a logistic regression analysis with repeated measures. The number of rating scores 2 and 3 (diagnostic and exemplary image quality) per technique and reader were chosen as characteristic values.
The side-by-side comparison ratings of overall quality of image pairs obtained in the same patient were assessed by descriptive statistics.

**Interobserver agreement**

As no standard of truth for the presence of disease was available, the interobserver agreement was used as a surrogate to evaluate the impact of image quality on diagnostic performance \(^{(12)}\). Interobserver agreement for the four groups of pathology was quantified for pairs of readers using weighted Cohen Kappa statistics. A $\kappa$ value of more than 0.4 was considered to represent moderate but clinically acceptable observer agreement, a kappa value of more than 0.6 was considered to represent good observer agreement, a value below 0.4 to represent only fair agreement \(^{(13)}\). The calculation was based on 5 reading categories (1 = definitely not present, 2 = probably not present, 3 = equivocal, 4 = probably present, 5 = definitely present).

**Results**

**Matching of patient groups**

A total of 150 bedside CXRs (n=50 CXR per acquisition technique) were obtained in 114 ICU patients. There was no significant difference between the three patient groups with respect to gender, age, BMI, patient entrance skin dose and effective dose (Table 1). Pulmonary consolidations were present in 13%, vascular congestion in 27%, pleural effusion in 37% and atelectasis in 54% of the patients. A significantly different prevalence was not found for any of the four diseases between the three different acquisition groups ($p = 0.32$, $p = 0.66$, $p = 0.12$ and $p = 0.65$, respectively; Table 2). None of the patients was suffering from a pneumothorax.

**Visibility of anatomical landmarks**

The landmarks located in the high attenuation area of the chest (trachea, carina, upper / lower thoracic spine and retro-cardiac vessels) were best seen with DR, followed by DR\(_{50}\%\); CR was the most inferior technique. Analysis of variance found significant differences between acquisition techniques ($p < 0.001$) and readers ($p < 0.001$), but no significant interactions ($P = 0.33$). The post hoc test found that all three techniques performed differently. For the remaining 4 landmarks located at density interfaces (bilateral diaphragmatic contours, azygo-esophageal recess and descending aorto-diaphragmatic recess) the post hoc test found a superiority of DR over DR\(_{50}\%\), but no significant difference between DR\(_{50}\%\) and CR.
Visibility of superimposed catheter fragments

The visibility of the catheter fragments superimposed over the upper abdomen was rated best with DR, followed by DR\(_{50\%}\); CR was the most inferior technique. An example is given in Figure 1. Analysis of variance found significant differences between techniques (p < 0.001) and readers (p = 0.001), but no significant interactions (p = 0.75). The post hoc test found that all three techniques performed differently.

**Table 1**

Characteristics of the patients of the three study groups.

<table>
<thead>
<tr>
<th>Acquisition technique</th>
<th>Age (years)</th>
<th>Gender (m/f)</th>
<th>BMI</th>
<th>mAs</th>
<th>Entrance skin dose (mGy)</th>
<th>Effective dose (mSv)</th>
</tr>
</thead>
<tbody>
<tr>
<td>CR</td>
<td>61</td>
<td>32/18</td>
<td>26.3</td>
<td>1.31</td>
<td>0.18</td>
<td>0.031</td>
</tr>
<tr>
<td>DR</td>
<td>59</td>
<td>30/20</td>
<td>26.3</td>
<td>1.16</td>
<td>0.16</td>
<td>0.029</td>
</tr>
<tr>
<td>DR(_{50%})</td>
<td>60</td>
<td>33/17</td>
<td>25.9</td>
<td>0.62</td>
<td>0.09</td>
<td>0.015</td>
</tr>
</tbody>
</table>

CR: computed radiography, DR: direct radiography, DR\(_{50\%}\): direct radiography at 50% dose reduction, BMI: body-mass index (weight / length\(^2\))

**Table 2**

Presence of disease in the three study groups.

<table>
<thead>
<tr>
<th></th>
<th>Parenchymal consolidations *</th>
<th>Vascular congestion</th>
<th>Pleural effusion</th>
<th>Atelectasis</th>
</tr>
</thead>
<tbody>
<tr>
<td>CR</td>
<td>10%</td>
<td>23%</td>
<td>48%</td>
<td>50%</td>
</tr>
<tr>
<td>DR(_{50%})</td>
<td>16%</td>
<td>32%</td>
<td>38%</td>
<td>54%</td>
</tr>
<tr>
<td>DR</td>
<td>12%</td>
<td>26%</td>
<td>36%</td>
<td>58%</td>
</tr>
<tr>
<td>Mean</td>
<td>13%</td>
<td>27%</td>
<td>37%</td>
<td>54%</td>
</tr>
</tbody>
</table>

=* Could be caused by a pneumonia, edema, ARDS, etc. Atelectasis was specifically excluded and rated separately. CR: computed radiography, DR: direct radiography, DR\(_{50\%}\): direct radiography at 50% dose reduction

**Figure 1**

Region of interest of bedside chest radiographs of one patient obtained on three different days with all three techniques. There is improved transparency of the mediastinum with DR and DR\(_{50\%}\) compared with CR with a superior delineation of the thoracic spine and the naso-gastric tube. Catheter fragments superimposed over the upper abdomen are best delineated with DR followed by DR\(_{50\%}\); CR is the most inferior technique.
Subjective assessment of overall image quality

The overall image quality of DR was significantly superior to CR (p = 0.006), there was no difference between overall image quality of DR$_{50\%}$ and CR (p = 0.85). Differences and significance of differences remained the same whether only scores 3 (exemplary image quality) or scores 2 and 3 (diagnostic and exemplary image quality) were included into the logistic regression analysis. Side-by-side comparison of DR versus CR was possible in 12 image pairs (48 ratings by four readers) and of DR$_{50\%}$ versus CR in 15 image pairs (60 ratings by four readers). Using the forced-decision model, the image quality of DR was ranked higher than CR in 96% (46/48) of ratings (Figure 2) and of DR$_{50\%}$ over CR in 87% (52/60) of ratings (Figure 3).

Interobserver agreement

With a κ-value of 0.48 the interobserver agreement with DR was moderate but clinically acceptable. The interobserver agreement with DR$_{50\%}$ was just below the threshold of clinically acceptable with a κ-value of 0.39; while CR was considerably below the threshold with a κ-value of 0.33 when averaged over all 4 readers. In general, averaged κ-values were fair for the assessment of consolidations (κ = 0.34) and vascular congestion (κ = 0.39) and moderate for the assessment of pleural effusions (κ = 0.57) and atelectasis (κ = 0.56).

Figure 2

CR versus DR images of the same patient in a side-by-side comparison. All four readers ranked the overall image quality of DR higher than that of the CR image. Note the superior delineation of the Swan-Ganz catheter and of the catheter fragments with DR.

Figure 3

CR and DR$_{50\%}$ images of the same patient as in Figure 2 in a side-by-side comparison. Although there is increased noise within the high-attenuation area of the mediastinum in the DR$_{50\%}$ image, the delineation of the Swan-Ganz catheter and of the catheter fragments is still superior with DR$_{50\%}$ as compared with CR. All four readers rated the image quality superior with DR$_{50\%}$. 
Discussion

Multiple studies have proven that dose reduction is feasible for upright chest radiography using flat-panel direct radiography (aSi TFT - DR) as compared with storage phosphor radiography (CR) without significant loss of image quality\(^{(1,4,5,14-16)}\). All clinical studies focussed on the delineation of anatomical landmarks and subjective scoring of image quality and uniformly report that DR, if obtained with the same dose as CR, provides superior image quality. When acquisition dose was reduced by a factor of two, DR achieved at least equivalent image quality compared with CR. Despite this dose reduction, DR still outperformed CR for some imaging aspects (Table 3). In order to being able to place these dose considerations into perspective, it is important to state not only relative but also absolute dose levels: reduction of a relatively high “standard dose” by a factor of 2 seems to be less of an achievement compared with operating on a generally lower dose level. Table 3 summarizes the clinical studies published recently for comparison of CR and DR. In only two of the six studies were absolute dose levels indicated: one referred to the entrance skin dose and the second to the detector dose. In the present study, the “standard dose” amounted to a mean of 0.18 mGy entrance skin dose for CR and DR and to 0.09 mGy entrance skin dose for the DR\(_{50\%}\). These dose levels are comparable to skin entrance dose levels published by Bacher et al. for upright chest radiography; 0.17 mGy for CR and 0.07 mGy for DR\(^{(3)}\). More than 10 years ago Leppek et al. had published two to three times higher entrance skin doses for conventional film-screen bedside chest radiography with a mean of 0.42 mGy (range 0.16-0.69 mGy)\(^{(16)}\). Although the single acquisition dose is rather low in bedside chest radiography and substantially reduced with the CR technique, as compared with conventional film-screen, dose reduction remains a relevant issue, since those patients often require multiple follow-up studies. Image quality requirements are different in bedside chest radiography compared with routine upright chest radiography. Contrast resolution has to be high, especially in high-attenuation areas of the mediastinum, to display the multiple types of catheter materials, whereas requirements for spatial resolution are generally low. The aim of the present study was to test whether dose reduction with DR is feasible for bedside chest radiographs of adults. To test the influence of dose reduction on image quality a number of reading measures, which included subjective scoring of visibility of anatomic landmarks and a relatively simple detection task, were applied. The latter referred to one of the major imaging tasks of bedside chest radiography, namely the assessment of monitor
### Table 3
List of clinical studies comparing direct radiography (DR) and computed radiography (CR) to assess the influence of potential dose reduction with DR.

<table>
<thead>
<tr>
<th>Author, journal and year of publication</th>
<th>n</th>
<th>Reference for comparison</th>
<th>Studied parameters</th>
<th>Results</th>
</tr>
</thead>
<tbody>
<tr>
<td>Goo</td>
<td>46</td>
<td>CR vs. SE-DR</td>
<td>11 anat. landmarks</td>
<td>Residents: DR &gt; CR (n=6 landmarks) CR &gt; DR (n=2 landmarks) Radiologist: DR &gt; CR (n=8 landmarks)</td>
</tr>
<tr>
<td>AJR 2000</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Biemans</td>
<td>63</td>
<td>CR vs. SE-DR</td>
<td>21 anat. landmarks</td>
<td>Postero-anterior: DR &gt; CR Lateral: DR = CR</td>
</tr>
<tr>
<td>Invest Rad 2002</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Ganten</td>
<td>30</td>
<td>CR vs. CsI-DR</td>
<td>10 anat. landmarks</td>
<td>DR&lt;50% = CR</td>
</tr>
<tr>
<td>AJR 2002</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Hermann</td>
<td>75</td>
<td>CR vs. CsI-DR</td>
<td>8 anat. landmarks</td>
<td>DR = DR&lt;50% = CR</td>
</tr>
<tr>
<td>Eur Rad 2002</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Bacher</td>
<td>100</td>
<td>CR vs. CsI-DR</td>
<td>Overall image quality</td>
<td>DR&lt;50% &gt; CR</td>
</tr>
<tr>
<td>AJR 2003</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Gruber</td>
<td>50</td>
<td>CR vs. DR</td>
<td>Overall image quality</td>
<td>Image quality: DR &gt; DR&lt;50% &gt; CR Landmarks in high attenuation area: DR and DR&lt;50% &gt; CR Landmarks in low attenuation area: DR &gt; CR; DR&lt;50% = CR</td>
</tr>
<tr>
<td>Eur Rad 2006</td>
<td></td>
<td></td>
<td>8 anat. landmarks</td>
<td>Visual perception of image noise</td>
</tr>
</tbody>
</table>

Indices (e.g., DR<50%) indicate the reduced level of acquisition dose in percentage of the “standard” dose. CR: computed radiography; SE-DR: flat-panel direct radiography using a selenium detector; CsI-DR: flat-panel direct radiograph using a Cesium-Iodine TFT detector.
devices, and was based on an absolute standard of truth. For all criteria that referred to image quality, DR was significantly superior to both DR\textsubscript{50\%} and CR. DR\textsubscript{50\%} was superior to CR for the delineation of catheter fragments and the visibility of anatomic landmarks in the high-attenuation area of the mediastinum, suggesting a strong superiority of DR over CR with respect to the signal-to-noise ratio in high-attenuation areas based on its higher dose efficiency. When overall image quality was evaluated separately per image, DR was found to be superior to CR whereas DR\textsubscript{50\%} and CR achieved equivalent image quality scores. During the side-by-side comparisons of image pairs using a forced-decision model, image qualities of both DR and DR\textsubscript{50\%} were rated superior to CR in 96\% and 87\%, respectively. These results support the fact that when judging image quality, a side-by-side comparison is far more sensitive in detecting subtle quality differences when compared with separate scoring of image quality per image. Regarding differences of image quality, in the absence of a pathological standard, we used interobserver agreement as a surrogate for diagnostic performance of the three different techniques. Four diseases frequently encountered at bedside chest radiography were tested. As no standard of truth was available, readers were not specifically asked to determine the underlying disease, but were only asked to assess the presence or absence of this type of imaging finding irrespective of severity and extent of findings. Only a fair to moderate interobserver agreement was found for the assessment of these four categories of intrapulmonary disease. It is common clinical experience that interobserver agreement is lower for bedside chest radiography than for upright radiography, which is most likely due to the sometimes vastly differing image quality and the non-specific nature of radiographic findings. It has to be noted that images were evaluated without the availability of previous radiographs and without any clinical context. As stated before, no specification of disease severity was required, which is likely to have caused some disagreement, especially for images with subtle findings. All of these aspects are likely to have contributed to the generally low interobserver agreement rates. In fact, the kappa values in the present study are probably lower than normally encountered under clinical conditions and are likely to underestimate the diagnostic potential of the bedside chest radiographs. However, under comparable (study) conditions DR was the only technique that on average achieved an agreement rate beyond the threshold of 0.4, which is considered as clinically acceptable. In that respect DR outperformed CR and also DR\textsubscript{50\%}. DR\textsubscript{50\%} also achieved at least the same agreement rate as CR. These results confirm previous experiences that DR either allows for improved image quality and potentially improved diagnostic
Discussion

performance if obtained with the same dose as that used for CR or alternatively allows for dose reduction on a slightly lower but unaltered diagnostic level comparable to CR. An advantage of the present study was that the image quality obtained via different methods was assessed, including subjective scoring of quality, grading of the delineation of anatomic features, and using visual tasks that are typical of the radiological evaluation methods of bedside chest radiographs of ICU patients, such as the delineation of catheter fragments and the assessment of the presence of disease. However, the present study suffers from the following limitations: first, the catheter fragments were relatively short and superimposed over the area of the upper abdomen and not within the mediastinum. Thus, position and length of the catheters lacked clinical characteristics; however, this technique allowed the visualisation of a range of monitoring devices under controlled conditions. Second, image pairs of the same patient obtained with two acquisition techniques were available in only a subgroup of the study patients. It should be noted requests for ICU bedside chest radiographs were strictly based on clinical indications; thus, patients did not undergo chest radiography daily. In addition, only a group of three technicians were involved in the acquisition of the study images with the advantage of minimizing external effects on image quality. As result, image acquisition was spread over a longer period of time. To minimize the effect of different patient groups, the study groups were statistically matched with respect to age, BMI, acquisition dose, and the presence of disease. Differences that may have potentially remained e.g., with respect to the extent of disease or the effects of different inspiration levels, cannot be excluded but are likely to be small as the choice of acquisition technique was randomly assigned. Third, acquisition parameters were set by the technicians according to the patient’s weight and the technician’s visual assessment of the patient’s constitution and condition as it is common practice in an ICU environment where an automatic exposure control is not available. Although it cannot be proved that every single DR image was acquired with exactly 50% less dose, on average DR images had been obtained either with the “standard” or with a dose lowered by 50%. From the present results it can be concluded that DR generally provides superior image quality to CR also at the bedside. A dose reduction of 50% can be applied without risking the loss of diagnostic information. At the same dose as CR, DR provides better delineation of landmarks and catheters and increases interobserver agreement on the assessment of disease categories.
References

12. Bankier AA, Levine D, Halpern EF, Kressel HY. Consensus interpretation in imaging research: is there a better way? Radiology 2010; 257:14-17