

Evaluation of low-energy contrast-enhanced spectral mammography images by comparing them to full-field digital mammography using EUREF image quality criteria

U. C. Lalji · C. R. L. P. N. Jeukens · I. Houben · P. J. Nelemans ·
R. E. van Engen · E. van Wylick · R. G. H. Beets-Tan ·
J. E. Wildberger · L. E. Paulis · M. B. I. Lobbes

Received: 24 December 2014 / Revised: 2 February 2015 / Accepted: 24 February 2015 / Published online: 27 March 2015
© The Author(s) 2015. This article is published with open access at Springerlink.com

Abstract

Objective Contrast-enhanced spectral mammography (CESM) examination results in a low-energy (LE) and contrast-enhanced image. The LE appears similar to a full-field digital mammogram (FFDM). Our aim was to evaluate LE CESM image quality by comparing it to FFDM using criteria defined by the European Reference Organization for Quality Assured Breast Screening and Diagnostic Services (EUREF).

Methods A total of 147 cases with both FFDM and LE images were independently scored by two experienced radiologists using these (20) EUREF criteria. Contrast detail measurements were performed using a dedicated phantom. Differences in image quality scores, average glandular dose, and contrast detail measurements between LE and FFDM were tested for statistical significance.

Results No significant differences in image quality scores were observed between LE and FFDM images for 17 out of

20 criteria. LE scored significantly lower on one criterion regarding the sharpness of the pectoral muscle ($p < 0.001$), and significantly better on two criteria on the visualization of micro-calcifications ($p = 0.02$ and $p = 0.034$). Dose and contrast detail measurements did not reveal any physical explanation for these observed differences.

Conclusions Low-energy CESM images are non-inferior to FFDM images. From this perspective FFDM can be omitted in patients with an indication for CESM.

Key Points

- Low-energy CESM images are non-inferior to FFDM images.
- Micro-calcifications are significantly more visible on LE CESM than on FFDM.
- There is no physical explanation for this improved visibility of micro-calcifications.
- There is no need for an extra FFDM when CESM is indicated.

U. C. Lalji · C. R. L. P. N. Jeukens · I. Houben · E. van Wylick ·
R. G. H. Beets-Tan · J. E. Wildberger · L. E. Paulis ·
M. B. I. Lobbes (✉)

Department of Radiology and Nuclear Medicine, Maastricht
University Medical Center, P.O. Box 5800, 6202
AZ Maastricht, The Netherlands
e-mail: marc.lobbes@mumc.nl

P. J. Nelemans
Department of Epidemiology, Maastricht University, P.O. Box 616,
6200 MD Maastricht, The Netherlands

R. E. van Engen
Dutch Reference Centre for Screening, Wijchenseweg 101,
P.O. Box 6873, 6503 GJ Nijmegen, The Netherlands

R. G. H. Beets-Tan · J. E. Wildberger · M. B. I. Lobbes
GROW School for Oncology and Developmental Biology,
Maastricht University Medical Centre, P.O. Box 616,
6200 Maastricht, The Netherlands

Keywords Breast cancer · Mammography · Quality control ·
CESM · CEDM

Abbreviations

CESM	Contrast-enhanced spectral mammography
LE	Low-energy
FFDM	Full-field digital mammography
DBCSP	Dutch Breast Cancer Screening Program
EUREF	European Reference Organization for Quality Assured Breast Screening and Diagnostic Services
ESAK	Entrance Surface Air Kerma
AGD	Average glandular dose
CBT	Compressed Breast Thickness

Introduction

In breast imaging, mammography continues to play an important role. The diagnostic performance of mammography is good, but depends heavily on breast density [1]. With the introduction of digital mammography, several applications that increase diagnostic accuracy (even in dense breasts) have been developed. One of these is contrast-enhanced spectral mammography (CESM). Recent studies continue to show that CESM is superior to conventional mammography in breast cancer detection [2–5].

A standard CESM examination consists of three sets of images per exposure: a low energy (LE) image, a high energy image, and a recombined image (Fig. 1). The LE image is acquired at peak kilovoltage (kVp) values ranging from 26–31 kVp. In this way, the entire x-ray spectrum is below the k-edge of iodine (which is 33.2 keV) [3]. As a result, the LE image resembles a full-field digital mammogram (FFDM), although iodine contrast is already present within the breast. However, in order to omit FFDM in cases for which CESM is indicated (for example in postmenopausal women with a highly suspicious palpable breast mass), LE images have to comply with the high image quality standards used in breast imaging and should at least be non-inferior to their FFDM counterpart. The high energy image is not used for diagnostic purposes, but is used in post-processing to create the recombined image, in which areas of enhancement can be appreciated.

In this retrospective study, we aimed to study the differences between LE and FFDM images in a standardized and structured manner. Our primary study aim was to evaluate the image quality of LE images using FFDM acquired at the Dutch Breast Cancer Screening Program (DBCSP) as a reference standard and using the image quality criteria defined by the European Reference Organization for Quality Assured

Breast Screening and Diagnostic Services (EUREF) [6]. Our secondary aim was to evaluate the differences in dose and contrast detail measurements between the two images.

Materials and methods

In The Netherlands, women aged 50–75 years are invited to participate in the DBCSP every 2 years [7]. If a breast abnormality is detected, women are referred to a hospital of their choice for further analysis. In our institute, every recalled patient is eligible to undergo CESM, unless contraindications for the use of iodine-based contrast agents exist, such as an increased risk of developing contrast-induced nephropathy or known contrast allergies.

In this study, data from women that underwent CESM for a screen-detected breast abnormality were collected between July 2013 and February 2014. Since they were recalled from screening, all women had a recent (screening) FFDM examination for comparison. The need for informed consent was waived by the local ethics committee.

Image acquisition and analysis

FFDM was performed on a standard mammography unit used in the DBSCP (Hologic Selenia, Bedford, MA, USA), whereas LE images were acquired as part of a CESM (Senographe* Essential with Senobright* upgrade, GE Healthcare, Chalfont St Giles, United Kingdom). All images were displayed on a dedicated mammography workstation (IDI MammoWorkstation 4.7.0, GE Healthcare, Chalfont St Giles, UK) equipped with mammography-approved monitors (Barco Coronis 5MP Mammo, Barco, Kortrijk, Belgium). Two experienced radiologists (U.L. and M.L., both viewing

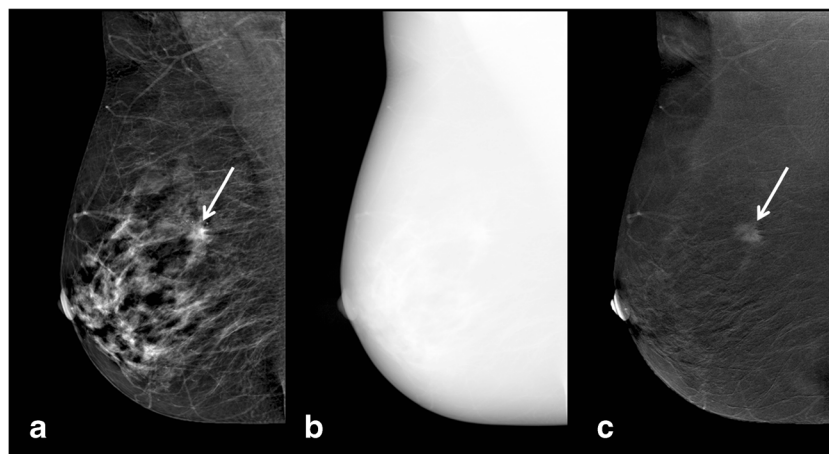


Fig. 1 A typical contrast-enhanced spectral mammography (CESM) examination (only right mediolateral oblique view shown), consisting of a low-energy (a), high-energy (b) and recombined (c) image. A suspicious lesion is seen on the low-energy image, showing

enhancement on the recombined image (white arrows). Histopathology showed invasive ductal carcinoma. The high-energy image is not for diagnostic purposes but is used for construction of the recombined image

over 6,000 clinical and screening mammographies annually) independently scored the images using 20 image quality criteria defined by EUREF (Table 1). These are criteria concerning overall diagnostic value, visibility of (micro-) calcifications, tissue visibility at the chest wall and at the skin line, overall density, as well as sharpness/noise impression [6]. The radiologists were not allowed to change window width or level settings while scoring the images.

The original EUREF criteria are dichotomous and can only be answered with ‘yes’ or ‘no’. In order to perform a more detailed analysis, we decided to use a five-point scoring scale, with ‘1’ being the least favourable score and ‘5’ the best possible score for each item. For example, for the question “Are the vascular structures visible through the dense parenchyma?” a score of ‘1’ would mean that vascular structures were hardly visible while a score of ‘5’ would indicate that visibility was excellent.

For scoring, the images were divided in two datasets, each containing either the FFDM or LE image of a patient. Patients were randomized with respect to the order in which images were evaluated: LE first and FFDM second, or vice versa. Randomization was performed to prevent recall bias. The

LE and FFDM images were viewed separately by both radiologists with a 6-week interval between the scoring of the first and second image set. Both radiologists were blinded for each other’s scores. In case of discrepancies consensus was reached by both radiologists favouring the lowest score given for that criterion.

Dose measurements

The Entrance Surface Air Kerma (ESAK) and Average Glandular Dose (AGD) were determined according to the method described in European guidelines [6]. In short, tube output of both mammography units was determined by measuring the air kerma using a dosimeter with a dedicated mammography probe calibrated for the target/filter combinations used in our measurements (Unfors RaySafe; Billdall, Sweden). The probe was placed on the bucky table below the compression paddle touching it.

Tube output was determined with current-time product 50 mAs (CESM)/70 mAs (FFDM) for all target/filter combinations observed in respectively the LE and FFDM patient groups. For each target/filter combination tube output

Table 1 Image quality criteria scores. The results of the maximum (5) score for FFDM and CESM in percentages and the median scores for FFDM and CESM are presented

Image quality criteria (<i>N</i> =147)	FFDM 5 score	LE 5 score	McNemar <i>p</i> -value	Median FFDM	Median CESM	Wilcoxon <i>p</i> -value
Is there a good visualization of the skin line?	93.2 %	97.3 %	0.109	5	5	0.058
Are the vascular structures visible through the dense parenchyma?	98.6 %	100.0 %	0.475	5	5	0.18
Is there a sharp visualization of the pectoral muscle?	40.8 %	21.1 %	<0.0001**	4	4	<0.0001**
Is there a good visualization of the Cooper’s ligaments and vascular structures in the subcutaneous and prepectoral area?	99.3 %	100.0 %	1	5	5	0.317
Are the micro-calcifications visualized and well outlined?*	95.0 %	100.0 %	0.042**	5	5	0.02**
Is there sufficient contrast in the dark areas?	99.3 %	98.6 %	1	5	5	0.564
Is there sufficient contrast in the white areas?	99.3 %	97.3 %	0.25	5	5	0.317
Is the glandular tissue sufficiently white?	100.0 %	99.3 %	1	5	5	0.317
Is the background sufficiently dark?	100.0 %	100.0 %	1	5	5	1
Do all images appear in the same way?	99.3 %	100.0 %	1	5	5	0.317
Is there disturbing noise in the dark areas?	100.0 %	100.0 %	1	5	5	1
Is there disturbing noise in the white areas?	99.3 %	99.3 %	1	5	5	0.317
Are there any artefacts?	99.3 %	100.0 %	1	5	5	0.317
Contrast in the white regions?	99.3 %	98.0 %	0.5	5	5	0.564
Contrast in the dark regions?	100.0 %	100.0 %	1	5	5	1
Overall contrast of the images?	99.3 %	98.6 %	1	5	5	0.317
Sharpness of the images?	100.0 %	100.0 %	1	5	5	1
How satisfied are you with the representation of micro-calcifications?*	95.3 %	100.0 %	0.074	5	5	0.034**
How satisfied are you with the representation of opacities?	100.0 %	99.3 %	1	5	5	0.317
How satisfied are you with the representation of the image?	97.3 %	96.6 %	1	5	5	0.655

* for *n*=120 cases

** statistically significant

FFDM full-field digital mammography, CESM contrast-enhanced spectral mammography

measurements of three tube voltages covering the whole clinically used range were measured. For the same spectra the half-value layer (HVL) was also measured using the dosimeter in the setup described in the European guidelines [6]. The technical parameters required for the ESAK and AGD calculation, i.e. kV, target, filter, current-time product and compressed breast thickness (CBT), were obtained for each exposure from the DICOM header of the images. The ESAK was calculated using the appropriate time-current product and tube output, which was corrected for breast thickness using the inverse square law. Next, the AGD was calculated from the ESAK using the Dance model [6, 8, 9].

Contrast detail phantom measurements

Contrast detail measurements were performed using a CDMAM phantom (version 3.4, Artinis Medical Systems, Elst, The Netherlands), which consists of a matrix of gold disks of diameter range 0.06–2.0 mm and thickness range 0.03–2.0 μm . The phantom was positioned in the middle of a stack of polymethylmethacrylate (PMMA) plates of varying thicknesses to simulate the clinical range of CBT as described in the EUREF guidelines [6], where the CDMAM phantom corresponds to 10 mm PMMA thickness. To determine the exposure parameters for each thickness, i.e. target/filter combination, tube voltage and current-time product, a FFDM or LE exposure was made in fully automatic exposure control mode. Using these exposure parameters, 16 exposures were made for each simulated breast thickness. The unprocessed images were exported to the analysis software (CDMAM analyzer, version 2.0.5, Elst, The Netherlands), which analyses the threshold detectability of the gold disks in each image, resulting in an average threshold thickness for each disc diameter. Subsequently, we plotted the threshold detection thickness versus the disc diameter on a log-log scale and generated a contrast-detail curve by performing a fit.

Statistical analyses

For each of the 20 EUREF criteria, the frequency of the (optimal) ‘5’ score was calculated. Any statistical significance between the scores of LE and FFDM images were assessed using a McNemar test for paired proportions. In addition, median scores of LE and FFDM images were calculated, using a Wilcoxon signed rank test to test for statistically significant differences. In case of micro-calcifications, if none were observed, ratings (criteria 5 and 18) were scored as ‘not applicable’.

For the dose measurements, AGD values of the entire patient population were tested pairwise for differences between FFDM and LE exposures using a Wilcoxon matched-pair signed rank test. Furthermore, we studied whether differences in AGD depended on breast thickness. To this extent, mean

CBT was calculated for each patient by averaging the CBT of all exposures taken on both systems, which were four FFDM and four LE exposures (CC and MLO orientation of left and right breast). Patients were categorized according to their mean CBT and the AGD was statistically tested for differences between the FFDM and LE groups using a paired sample *t*-test.

To compare the detection curves between FFDM and LE in the contrast detail measurements a multivariable linear regression analysis was performed to evaluate detection threshold thickness differences for LE and FFDM. The mammography method (FFDM or LE) and 1/diameter of the gold discs were used as covariates, the latter as a first-order approximation of the theoretical relation stated in the European guidelines [4].

All *p*-values <0.05 were considered to be statistically significant. All statistical analyses were performed using SPSS (IBM SPSS Statistics for Windows, Version 22.0. Armonk, NY, USA).

Results

A total of 147 women who underwent CESM for a screen-detected breast abnormality were included. A total of 588 FFDM images were compared to 588 LE images. Mean age was 59.5 years (range 49–75 years). Thirty-six of 147 women were diagnosed with breast cancer (24.5 %). Of these women, seven were diagnosed with DCIS (4.8 %) and the remaining 29 with invasive carcinoma (19.7 %). The remaining women were diagnosed with a benign condition as reason for their referral (such as superposition of glandular tissue as cause for the abnormality, cysts, fibro-adenoma, papilloma, etc.). The mean time between FFDM at time of referral and CESM was 14 days (SD 11 days).

The results of the scores are summarized in Table 1. For FFDM images, the frequency of the maximum (‘5’) score per criterion ranged from 40.8 % to 100 %. For LE, this ranged from 20.1 % to 100 %. For 19 criteria the median score for both FFDM and LE was 5 with an exception of the visibility of the pectoral muscle, where the median score of both FFDM and CESM was 4. In 17 of the 20 image quality criteria, no significant differences between FFDM and LE were observed. One criterion concerning the visibility of the pectoral muscle showed a maximal score frequency of 40.8 % for FFDM and 20.1 % for LE ($p < 0.0001$). Of the two remaining criteria (both regarding visualization of micro-calcifications) LE score was better than FFDM score ($p = 0.02$ and $p = 0.034$, respectively; Table 1).

Figure 2 presents the AGD values of FFDM and LE exposures as a function of compressed breast thickness of all patients. The average pairwise difference of the AGD values was 42 %, where the AGD of the LE was higher than that of FFDM (Table 2). Pairwise analysis showed that not only for

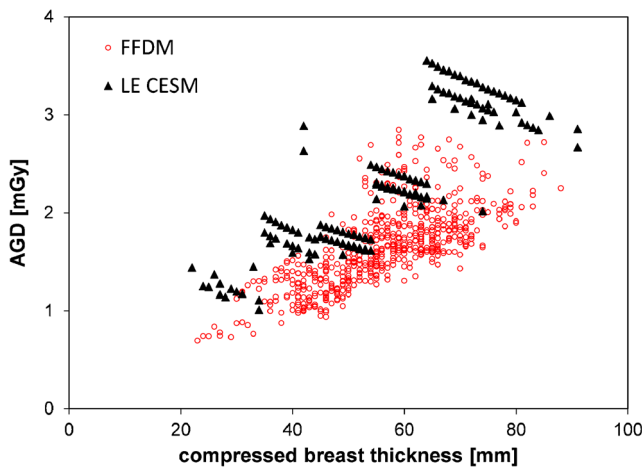


Fig. 2 Average glandular dose (AGD) of the complete patient group values plotted against compressed breast thickness for the full-field digital mammography (FFDM) (open red circles) and low-energy (LE) (solid black triangles) exposures

the whole patient population but also for all breast thickness categories, the LE AGD values were significantly higher than the FFDM AGD values (Table 2 and 3). Both LE CESM and FFDM AGD-values were in compliance with the acceptable limits of the EUREF guidelines [6].

The settings used in the CDMAM phantom measurements (Table 2) and the resulting AGD values (compare Table 2 to Fig. 2) are representative for settings used in the patient groups. Furthermore, the AGD values used in all measurements are well below the achievable dose levels reported in the EUREF guidelines [6], except for one that is well below the acceptable limit. Figure 3 shows the threshold thickness of the observed gold disks versus the gold disk diameter, for each phantom thickness. The threshold contrast visibility for both systems is well below the achievable limit values for a standard breast thickness of 60 mm (50 mm PMMA equivalent) stated in the EUREF guidelines [6]. For all breast thicknesses no significant differences in detection threshold thickness were found between the FFDM and LE curves (*p*-values of

0.29, 0.10, 0.13, 0.42 and 0.12 for breast thickness of 30, 45, 60, 75 and 90 mm, respectively).

Discussion

CESM is a promising new breast imaging modality that has been shown to be superior to conventional mammography in the detection of breast cancer [2, 3, 10]. No studies so far have compared the quality of the LE images to that of FFDM in a standardized and structured manner, using well established EUREF image quality criteria. We demonstrated that for the majority of image quality criteria there were neither clinically relevant nor statistically significant differences between these two modalities. Interestingly, we observed that the visualization of micro-calcifications is superior in LE images. Based on additional dose and contrast detail measurements, we could not find any physical explanation for these observations.

In 17 of 20 EUREF criteria no statistically significant differences were observed, indicating that there is hardly any difference between FFDM and LE in terms of image quality. The delineation of the pectoral muscle was scored significantly lower for LE when compared to FFDM. However, in our experience this will not cause any relevant diagnostic dilemmas, as the recombined image can also be used for assessment of lesions covering the pectoral muscle. In a worst-case scenario, an additional FFDM in mediolateral oblique view can still be considered if the lesion of interest is adjacent to the pectoral muscle. Figure 4 shows an example of the obscured pectoral muscle on LE versus FFDM. More interestingly, we observed a statistically significant improvement of micro-calcification visualization in LE images when compared to FFDM. This latter finding can indeed be of interest, as it suggests that ductal carcinoma in situ can be diagnosed more confidently or its extent might be assessed more accurately using LE CESM images. However, the limited number of DCIS cases (*n*=7) in this study restrained us from further testing this hypothesis.

Table 2 Phantom composition and settings used in the CDMAM measurements

Equivalent breast thickness (mm)	Total PMMA thickness (mm)	Phantom composition* (mm)	FFDM					LE CESM				
			Target	Filter	Tube voltage (kV)	Tube load (mAs)	AGD (mGy)	Target	Filter	Tube voltage (kV)	Tube load (mAs)	AGD (mGy)
32	30	10/Ph/10	W	Rh	26	61	0.79	Mo	Rh	27	45	1.1
45	40	10/Ph/10/10	W	Rh	28	86	1.1	Rh	Rh	29	71	1.8
60	50	10/10/Ph/10/10	W	Rh	32	81	1.3	Rh	Rh	31	80	2.3
75	60	25/Ph/25	W	Ag	30	111	1.7	Rh	Rh	30	140	3.6
90	70	25/Ph/25/10	W	Ag	34	96	1.8	Rh	Rh	30	140	3.8

*Ph=CDMAM phantom, 10 and 25 refer to the thickness of the PMMA plates (mm)

FFDM full-field digital mammography, LE low energy, CESM contrast-enhanced spectral mammography, AGD average glandular dose

Table 3 AGD per exposure for FFDM and LE images for five compressed breast thickness categories and for all patients, *p*-values denote the significance of a pairwise comparison

Compressed breast thickness (mm)	Number of patients	FFDM AGD mean (mGy)	± SD (mGy)	Range (mGy)	LE CESM mean (mGy)	± SD (mGy)	Range (mGy)	Significance
<40 mm	13	1.15	0.24	0.69–1.75	1.6	0.3	1.0 – 2.0	<0.001
41–50 mm	31	1.3	0.3	0.9–2.6	1.8	0.1	1.5 – 2.5	<0.001
51–70 mm	87	1.8	0.3	1.1–2.8	2.5	0.6	1.6 – 3.6	<0.001
71–80 mm	13	2.0	0.3	1.4–2.7	3.22	0.12	3.03 – 3.52	<0.001
>81 mm	3	2.19	0.23	1.90–2.71	2.95	0.18	2.67 – 3.26	<0.001
All	147	1.7	0.4	0.7–2.8	2.3	0.7	1.0 – 3.6	<0.001

FFDM full-field digital mammography, CESM contrast-enhanced spectral mammography, AGD average glandular dose, LE low energy

To further explain the differences in image quality scores we performed additional dose and contrast detail measurements. The average AGD of the FFDM exposure is 1.7 ± 0.4 mGy, while the AGD of the LE exposure is 2.3 ± 0.7 mGy. The difference reflects differences in AOP settings

for FFDM and LE CESM, i.e. target/filter combination, tube voltage and tube load, where for LE CESM the prerequisite was that the LE acquisition should be suitable for reconstruction of the contrast-enhanced image. Furthermore, the doses used in these examinations were in agreement with results

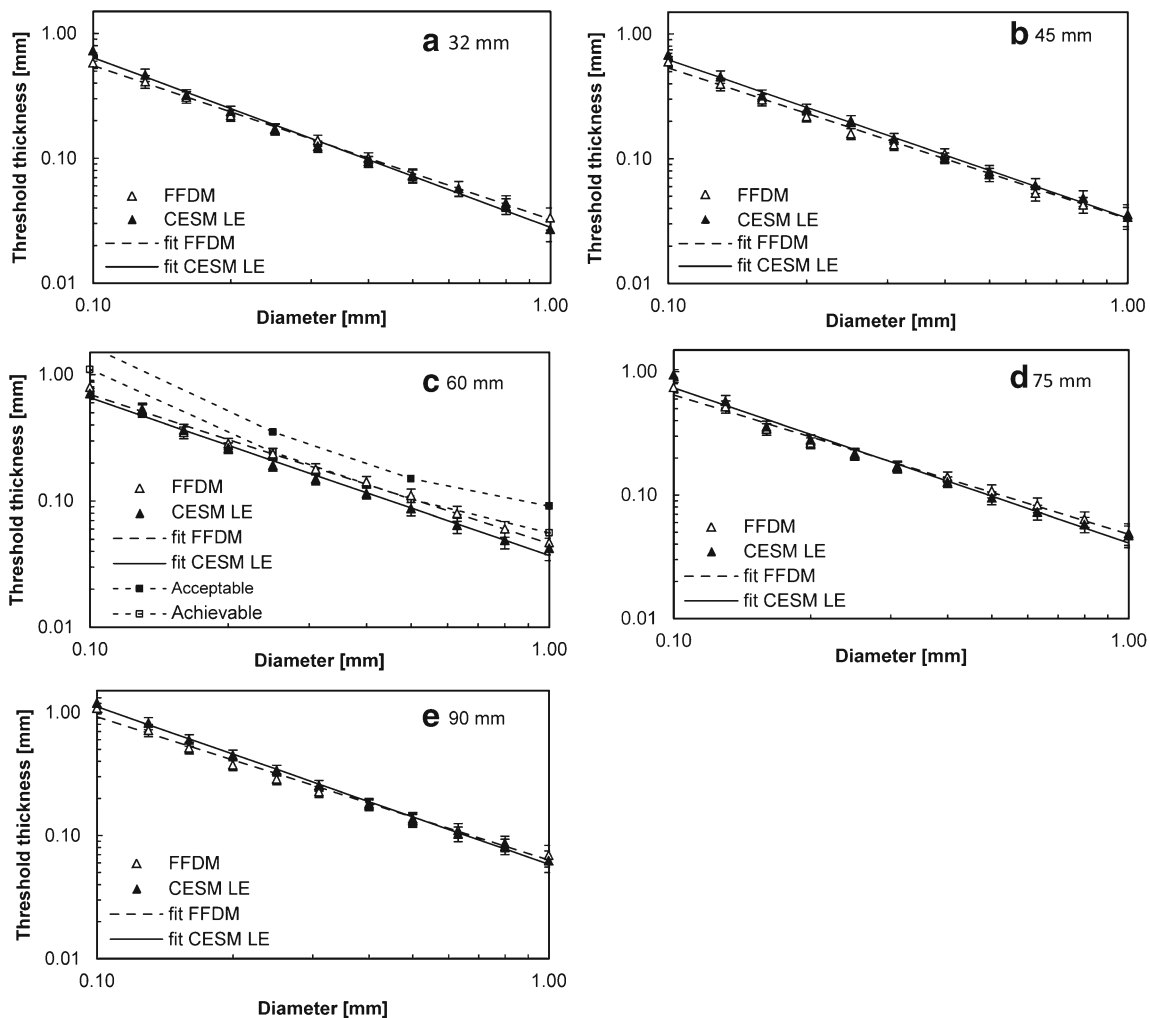


Fig. 3 Threshold gold thickness detected as a function of gold disk diameter for a range of phantom thicknesses corresponding to breast thicknesses from (A) 32 mm to (E) 90 mm. The solid and dashed lines

are data fits. In (C) the acceptable and achievable limits of the EUREF guidelines are also shown

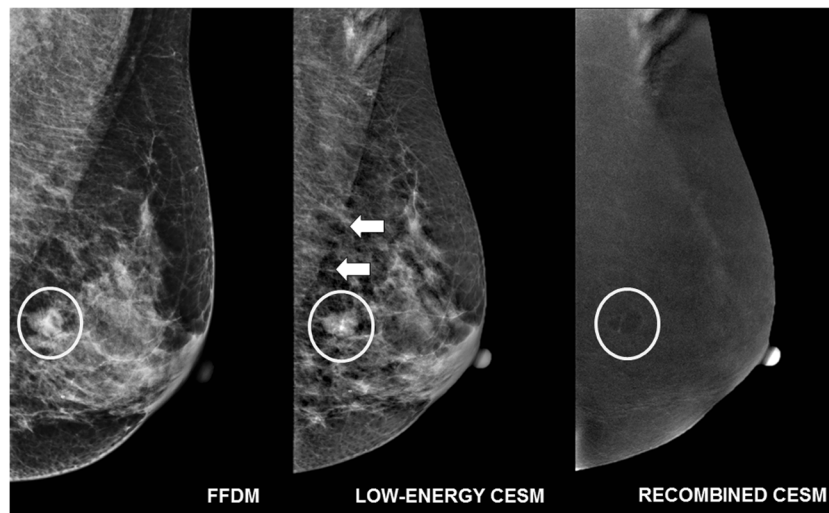


Fig. 4 A 54-year-old female recalled from the breast cancer screening program (full-field digital mammography (FFDM) image) for a round mass in the left breast (*), also visible on the low-energy contrast-enhanced spectral mammography (CESM) image. On the recombined image, an ‘eclipse sign’ is visible, suggesting a cyst, and confirmed by

targeted ultrasound. In this case, the sharpness of the pectoral muscle was scored ‘5’ on the FFDM image. On the low-energy CESM image, the delineation of the pectoral muscle was lost (white arrows), resulting in a score of ‘3’ from both radiologists

from a previous study [11] showing that the observed image quality differences could not be explained by using abnormal radiation doses in one of the techniques. We also performed CDMAM contrast detail measurements to investigate whether the observed differences in scoring could be explained. The different scores concerning pectoral muscle visibility could not be assessed by these measurements as this structure is not represented in the phantom. The detectability of smaller diameter gold disks in this phantom was found to relate to micro-calcification detection [12]. We did not observe a significant difference in the contrast detail curves, indicating that FFDM and LE CESM have similar contrast detail detection. FFDM and LE CESM both comply with the acceptable as well as the achievable limits for contrast threshold visibility. Since the CDMAM analysis was performed on the unprocessed images, it can be concluded that the observed differences were not due to hardware differences. A possible explanation for the improved calcification detection may be found in different post-processing techniques performed to obtain the final image from the unprocessed images. Unfortunately, we are not able to further evaluate the effect of these post-processing steps in our phantom or clinical measurements.

Only two previous papers focused on the potential of LE images as replacement for FFDM. Fallenberg et al. [13] focused on the ability of cancer detection and size estimation of bilateral CESM, digital mammography, and the combination of CESM and mammography in 107 patients. They did not use image quality criteria such as those suggested by EUREF. In their study, sensitivity of mammography was 77.9 %, which increased to approximately 95 % when using CESM alone or the combination of CESM and mammography, while CESM alone showed the best correlation with pathology

regarding tumour size measurements. Although there are minimal and hardly noticeable differences in attenuation due to the iodine contrast agent already present in the breast during the LE acquisition, the improved sensitivity was suggested to be attributed to a slight enhancement of cancers even on the LE images. In addition, our observation suggests that micro-calcifications can be better visualized on LE images. Most importantly, neither Fallenberg et al.’s study nor our study showed any signs of inferiority of LE images compared to mammography.

Francescone et al. [14] studied 170 examinations in 88 women and compared the LE images with FFDM images that were obtained within 6 months. They mainly focused on technical parameters including posterior nipple line distance, compression thickness and compression force on the MLO projection. Similar to Fallenberg et al., they compared tumour size measurements on both modalities. An important limitation to this study was that the readers were not blinded and analysis was performed in a side-by-side comparison, thereby potentially introducing bias. No statistically significant differences were observed for all studies’ parameters between LE and FFDM, but this study also did not use standardized image quality criteria such as suggested by EUREF.

Although the studies of Fallenberg et al. and Francescone et al. suggested non-inferiority of LE CESM compared to FFDM, our study contains important additional information as it assessed differences by using well-defined image quality criteria in a structured and standardized manner in the largest study population so far. Nevertheless, when all current evidence is summarized, it is valid and safe to conclude that LE CESM images are indeed non-inferior to FFDM, and might even be superior in terms of breast cancer detection and micro-

calcification evaluation. This means that FFDM can be omitted if a direct indication for CESM exists, for example in (postmenopausal) women with a suspicious palpable breast mass, or women with incidental suspicious breast findings observed in other imaging exams, such as chest CT examinations.

Our study has some limitations. First, FFDM images were made with a different mammography unit than the LE images. Both radiologists, however, were experienced in reading images from both mammography units. Consequently, they could distinguish the FFDM from the LE images based on their appearance alone. However, our current (random) reading strategy, including a 6-week interval, minimized the effect of any recall bias. Another limitation is that in our contrast detail phantom studies, we used the unprocessed images for data analysis. This is the standardized method, recommended by EUREF, as the post-processing may introduce image artefacts and may slightly differ from the mammography post-processing as an object-dependent histogram of texture analysis may be used. Therefore, it was not possible to determine whether post-processing improved the contrast-detail detection, and, thus, whether this would be an explanation for the significantly improved detection of micro-calcifications by LE images.

Conclusion

The LE images obtained as part of a CESM examination are similar to FFDM for the vast majority of EUREF image quality criteria. The slightly poorer delineation of the pectoral muscle on LE images does not seem to cause diagnostic dilemmas. More importantly, micro-calcifications are better visualized on LE images when compared to FFDM. In conclusion, using LE images for diagnostic purposes is safe.

Acknowledgments The scientific guarantor of this publication is Dr. M.B.I. Lobbes, M.D., PhD. The authors of this manuscript declare no relationships with any companies whose products or services may be related to the subject matter of the article. The authors state that this work has not received any funding. P.J. Nelemans, PhD. kindly provided statistical advice for this manuscript. Institutional Review Board approval was obtained.

Written informed consent was waived by the Institutional Review Board. Methodology: retrospective, observational, performed at one institution.

Open Access This article is distributed under the terms of the Creative Commons Attribution Noncommercial License which permits any non-commercial use, distribution, and reproduction in any medium, provided the original author(s) and the source are credited.

References

1. Carney PA, Miglioretti DL, Yankaskas BC et al (2003) Individual and combined effects of age, breast density, and hormone replacement therapy use on the accuracy of screening mammography. *Ann Intern Med* 138:168–175
2. Lalji U, Lobbes M (2014) Contrast-enhanced dual-energy mammography: a promising new imaging tool in breast cancer detection. *Womens Health (Lond Engl)* 10:289–298
3. Lobbes MBI, Smidt ML, Houwers J et al (2013) Contrast enhanced mammography: techniques, current results, and potential indications. *Clin Radiol* 68:935–944
4. Blum KS, Rubbert C, Mathys B et al (2014) Use of contrast-enhanced spectral mammography for intramammary cancer staging: preliminary results. *Acad Radiol* 21:1363–1369
5. Luczyńska E, Heinze-Paluchowska S, Dyczek S et al (2014) Contrast-enhanced spectral mammography: comparison with conventional mammography and histopathology in 152 women. *Korean J Radiol* 15:689–696
6. EUREF European Guidelines - EUREF | European Reference Organisation for Quality Assured Breast Screening and Diagnostic Services (2013) European guidelines for Quality Assurance in breast cancer screening and diagnosis, Fourth edition Supplements. available at: <http://www.euref.org/european-guidelines>
7. Timmers JMH, Verbeek ALM, IntHout J, et al. (2013) Breast cancer risk prediction model: a nomogram based on common mammographic screening findings. *Eur Radiol* 23:2413–9. doi: 10.1007/s00330-013-2836-88
8. Dance DR, Skinner CL, Young KC et al (2000) Additional factors for the estimation of mean glandular breast dose using the UK mammography dosimetry protocol. *Phys Med Biol* 45: 3225–3240
9. Dance DR, Young KC, van Engen RE (2009) Further factors for the estimation of mean glandular dose using the United Kingdom, European and IAEA breast dosimetry protocols. *Phys Med Biol* 54: 4361–4372
10. Lobbes MBI, Lalji U, Houwers J et al (2014) Contrast-enhanced spectral mammography in patients referred from the breast cancer screening programme. *Eur Radiol* 24:1668–1676
11. Jeukens CRLPN, Lalji UC, Meijer E et al (2014) Radiation exposure of contrast-enhanced spectral mammography compared with full-field digital mammography. *Invest Radiol*. doi:10.1097/RLI.0000000000000068
12. Warren LM, Green FH, Shrestha L et al (2013) Validation of simulation of calcifications for observer studies in digital mammography. *Phys Med Biol* 58:N217–N228
13. Fallenberg EM, Dromain C, Diekmann F et al (2014) Contrast-enhanced spectral mammography: does mammography provide additional clinical benefits or can some radiation exposure be avoided? *Breast Cancer Res Treat* 146:371–381
14. Francescone MA, Jochelson MS, Dershaw DD et al (2014) Low energy mammogram obtained in contrast-enhanced digital mammography (CEDM) is comparable to routine full-field digital mammography (FFDM). *Eur J Radiol* 83:1350–1355