



Comparison of Imaging Quality between 2D Synthesised Mammograms Reconstructed from Digital Breast Tomosynthesis and 2D Full-field Digital Mammograms

Johannes Peters^{1*}, Gudrun Peters², Debra Ardath Meerkotter², Katie Daniels³
and Jonathan Shulman²

¹Medical School, University of Tasmania, Australia.

²Regional Imaging, Tasmania, I-MED Network, Australia.

³Royal Hobart Hospital, Australia.

Authors' contributions

This work was carried out in collaboration between all authors. Authors JP and GP designed the study, wrote the protocol and wrote the first draft of the manuscript. Author JP managed the literature searches and analyses of the study. Author GP oversaw the test set reading. All authors read, edited and approved the final manuscript.

Article Information

DOI: 10.9734/BJMMR/2016/29709

Editor(s):

(1) Arun Kumar Nalla, College of Medicine, University of Illinois, Peoria, IL, USA.

(2) Alex Xiucheng Fan, Department of Biochemistry and Molecular Biology, University of Florida, USA.

Reviewers:

(1) Roland Sennerstam, Karolinska Hospital and Karolinska Institutet, Sweden.

(2) Anonymous, Pennsylvania State University, USA.

(3) Michal Strzelecki, Lodz University of Technology, Poland.

Complete Peer review History: <http://www.sciencedomain.org/review-history/16893>

Original Research Article

Received 25th September 2016

Accepted 8th November 2016

Published 11th November 2016

ABSTRACT

Objectives: The aim of this study was to evaluate if 2D synthesised mammograms reconstructed from a digital breast tomosynthesis 3D data set are noninferior in imaging quality when compared to 2D full field digital mammograms.

Methods: A random sample of 100 mammograms was selected from a dedicated private breast imaging service in Australia. Selected cases were classified as normal, benign or malignant. Five breast radiologists retrospectively rated the overall imaging quality and lesion quality of 2D synthesised mammograms when compared with 2D full field digital mammograms. Cases with

*Corresponding author: E-mail: johannesp99@gmail.com;

metal artefact were reassessed using metal artefact post processing software.

Results: Overall image quality for all cases before metal reduction post processing was 0.39 (CI 0.32, 0.45) and after post processing 0.40 (CI 0.33, 0.46). Overall lesion quality for all cases before metal reduction post processing was 0.46 (CI 0.28, 0.65) and after post processing 0.58 (CI 0.43, 0.72). Results confirm noninferiority of both overall image quality and overall lesion quality when comparing 2D synthesised mammograms to 2D full field digital mammograms. Metal artefact reduction had an impact on improving ratings for 2D synthesised mammograms.

Conclusions: 2D synthesised mammograms reconstructed from a digital breast tomosynthesis 3D data set are noninferior when compared to 2D full field digital mammograms. This results in reduction of radiation dose and time under compression.

Advances in Knowledge: 2D synthesised mammograms reconstructed from a digital breast tomosynthesis 3D data set can replace 2D full field digital mammograms. Metal artefact reduction software should be used routinely.

Keywords: Tomosynthesis; mammography; synthesised mammograms; noninferiority.

1. INTRODUCTION

Mammography has been a widely accepted method for screening and early detection of breast cancer for many years [1]. Sensitivity for both analogue and digital mammography are reported to be between 36-70% depending on breast density [2]. Due to the limitations of mammography, particularly with regards to tissue superimpositions which can hide or mimic pathologies leading to reduced sensitivity and increased false-positives, new modalities have been sought to replace it [3,4]. A new imaging modality has emerged in recent years: Digital breast tomosynthesis (DBT) but only with the development of fast reading digital detectors has it become practical [5-7].

DBT uses x-rays and a digital detector where a limited arc of the x-ray tube collects a series of low dose cross sectional slices which are digitally reconstructed into approximately 1mm thick slices. By constructing this 3D volume into thin slices, there is a reduction in the degrading effect of superimposed tissue [4-6,8-10].

Data indicates that the use of DBT as an adjunct to full field digital mammography (FFDM) can improve cancer detection and reduce false-positive mammography results [6-9].

Currently DBT is used in adjunction to FFDM [11-13] to allow direct comparison as a standard practice of current 2D images with prior 2D images. The one primary concern is the discrepancies in detectability parameters between the two modalities [6,9,11]. Moreover segmental and clustered microcalcifications

are more easily and quickly appreciated with 2D mammographic images because microcalcifications can traverse multiple slices in 3D image [6,9].

By supplementing FDDM with DBT there is an increase in radiation dose [12]. This combined FDDM & DBT dose is approximately double the radiation dose of normal FFDM [11] although it still remains within acceptable levels [6]. The combined procedure also increases acquisition time [6,13].

A very recent development in software for DBT (C-View, Hologic, Bedford, Conn) has made it possible to reconstruct a 2D mammographic image, a synthesised mammogram (SM), from the 3D dataset. This would negate the need to pair a FFDM with the DBT and ultimately reduce the radiation dose by half [12]. So far only very limited published data [11,13] is available to explore whether 2D synthesised mammographic images are of same imaging quality as true digital mammography images. Zuley et al. [13] found SM alone or in combination with tomosynthesis is comparable in performance to FFDM alone. Very recent clinical results have been very encouraging saying that 2D synthesised mammograms are good enough to replace FFDM [14].

This study will explore whether 2D synthesised mammographic images are non-inferior to true digital mammography images as the current standard test. Ultimately the results may be used to determine if digital mammography can be replaced by synthesised images leading to reduction of radiation exposure and time under compression for the patient.

2. MATERIALS AND METHODS

2.1 Patients

100 patients were selected from the database of a private practice offering a dedicated diagnostic breast imaging service including DBT, ultrasound and breast MRI. Patients attending the private practise are either symptomatic or require follow up for a previous history of breast cancer. They are either referred by their primary care doctor or a specialist (breast surgeon, oncologist or radiation oncologist). Patients have to wait 5 years following their breast cancer diagnosis until they are eligible again for the public screening program in Australia.

The 100 cases were selected within a 3 month timeframe to ensure a representative sample of the patient demographics at the diagnostic breast service. The 100 patients assessed in the study were the only ones that met the inclusion criteria described in the method over that 3 month period.

Inclusion criteria for this study were (a) bilateral digital mammography images and DBT with 2D synthesised mammograms and (b) biopsy and pathology for malignant cases.

The selected cases were classified as 'normal', 'benign' or 'malignant' in the research data set. 'Normal' cases had breast ultrasound (uni- or bilateral depending on breast density) and/or no reported interval change when compared to prior imaging.

'Benign' cases had breast ultrasound (uni- or bilateral depending on breast density), comparisons with prior imaging and shown no interval change over a 12 month period +/- benign results if biopsies were undertaken.

'Malignant' cases were biopsy proven cancers.

Additionally breast density grading according to BI-RADs classification [15], size of malignant lesions and prior history were extracted from existing patient data.

The images for the study were acquired using the Hologic Selenia Dimensions AWS 8000 System (Hologic, Bedford, Conn) using the ComboHD mode which captures both a digital mammogram and tomosynthesis data set. The device uses high power tungsten anode with x-ray filters made of rhodium (Rh), silver (Ag) and

aluminium (Al). The image detector element size is 70 microns.

The X Ray tube moves across a 15 degree arc acquiring a series of low dose projections and digitally reconstructing them into 1mm thick slices. The radiation dose for an ACR phantom was 1.2 mGy for 2D digital mammography, 1.45 mGy for 3D DBT and 2.65mGy the ComboHD mode (2D + 3D mammography). Additionally the image processing software C-View (Hologic, Bedford, Conn) was used to produce a synthetic mammogram from this 3D dataset. These synthetic mammograms were compared to standard digital mammograms in the study.

The radiologist participating in the study viewed the images on a Hologic SecurViewDX 400 Workstation (Hologic, Bedford, Conn) using a 5 megapixel liquid crystal display (LCD).

5 Breast Imaging Radiologists ranging from 5 years to 18 years (mean=11.2 years) experience in breast imaging participated in this study. For this study the radiologists were asked to retrospectively rate the subjective imaging quality of 2D synthesised mammograms when compared to digital mammography as the standard test. A specific hanging protocol was created for the study which enabled radiologists to compare digital mammography craniocaudal (CC) and mediolateral oblique (MLO) views side-by-side with 2D synthesised mammograms of the same projections.

A rating system was developed to compare the image quality of 2D synthesised mammograms with digital mammograms. The 4 ratings were that 2D synthesised mammograms were 'same', 'better', 'worse but acceptable' or 'worse' than the digital mammograms.

Radiologists rated overall impression of the imaging quality of 2D synthesised mammograms when compared to digital mammograms for all cases ('normal', 'benign' and 'malignant') using the rating system.

Additionally for cases with appreciable lesions ('benign' and 'malignant') the radiologists were asked to rate the overall impression of the imaging quality of the lesion of 2D synthesised mammograms when compared to digital mammograms using the rating system.

Metal artefact cases were reassessed using metal artefact post processing software 'De-

metal' (Hologic, Bedford, Conn). The metal artefact software reduces flare on reconstructed/tomosynthesis images caused by high density objects in the breast such as biopsy clips. A button is essentially clicked when required and the new algorithm applied. This can be done after the image was acquired. The user can then accept or undo the new algorithm. This software was utilised, as it is the software made available to our practise by Hologic for the Hologic Selenia Dimensions AWS 8000 System (Hologic, Bedford, Conn) which was used to capture and process the mammograms. The same rating system as for all previous cases was used.

3. RESULTS

3.1 Description Dataset

The test set was composed of 100 patients ranging in age from 36 to 82 years (median, 59 years). 40 of the 100 cases were classified as 'normal'. Eight were classified as 'malignant' this included 5 mass lesions, 1 architectural distortion, 1 stellate lesion and 1 asymmetrical density. Mean size of malignant lesions was 26.6 mm (range: 12-45 mm) in maximum diameter. 52 cases were classified as benign this included 38 surgical scars, 8 cysts, 1 oil cyst, 1 radial scar, 1 macrocalcification, 1 foreign body reaction and 2 patients had both a cyst and a scar. 42 patients had a previous history of surgery (38 previous history of breast cancer and 4 benign surgeries).

19 cases (19%) were graded as breast density of 1, 32 cases (32%) a breast density of 2, 32 cases (32%) a breast density of 3 and 17 cases (17%) a breast density of 4.

3.2 Data Analysis

The ratings given by the five radiologists ('same', 'better', 'worse but acceptable' or 'worse' than the digital mammograms) were assigned numerical values "-2 worse", "-1 worse but acceptable", "0 same" and "2 better" to allow for calculation of inter-reader reliability and to calculate the 90% confidence interval and show non-inferiority.

The inter-reader agreement was calculated using Cohen's Kappa and Fleiss Kappa. Cohen's Kappa coefficient is a statistic which allows for the measurement of inter-reader agreement (reliability) for qualitative items. This can be used for two different ratings of the same reader or for a single rating from two different readers. It is a

more accurate measure than a simple percentage agreement calculation, as it takes into account that an inter-reader agreement could occur due to chance. Fleiss Kappa allows the inter reader agreement to be calculated between all the readers for an item.

The study was aimed to determine whether there is noninferiority between 2D synthesised mammograms and digital mammograms rather than superiority. Superiority is generally shown by demonstrating there is significant difference ($P \leq 0.05$) in scores between groups using a p-value. In this study, to determine non inferiority, a two sided 90% Wald confidence interval was calculated using generalised estimating equations which allow ratings from all 5 radiologists to be factored into the calculations (each case was rated by 5 radiologists). This procedure was also used to analyse the influence of breast density.

IBM SPSS 21 was used for the statistical analysis using generalised estimating equations. A linear-link function was utilised.

If the 90% confidence interval includes values ≤ 0 , the mean result is not significantly ($P > 0.05$) above 0 and non-inferiority has not been shown.

If all the values of the 90% confidence interval are > 0 it has the same meaning as $P \leq 0.05$ and would therefore show that this result shows non inferiority.

The mean score for all 5 readers in rating the **overall image quality** for all cases before metal reduction post processing was 0.39 (CI 0.32, 0.45) and after post processing 0.40 (CI 0.33, 0.46) (Table 1).

The mean score for all 5 readers in rating the **overall lesion quality for all cases** before metal reduction post processing was 0.46 (CI 0.28, 0.65) and after post processing 0.58 (CI 0.43, 0.72) (Table 2).

The mean score for all 5 readers in rating the **benign lesion** group on the **overall lesion quality** before metal reduction post processing was 0.41 (CI 0.20, 0.61) and after post processing 0.54 (CI 0.39, 0.70).

The mean score for all 5 readers in rating the benign lesion 'scar' sub-group on the **overall lesion quality** before metal reduction post processing was 0.43 (CI 0.17, 0.70) and after post processing 0.63 (CI 0.45, 0.81).

Table 1. Overall image quality after image editing

Radiologist	Mean	Std. error	90% Wald confidence interval	
			Lower	Upper
1	.16	.055	.07	.25
2	.52	.088	.38	.66
3	1.16	.099	1.00	1.32
4	.20	.073	.08	.32
5	-.06	.024	-.10	-.02
Total	.40	.039	.33	.46

Table 2. Overall lesion quality after image editing

Radiologist	Mean	Std. error	90% Wald confidence interval	
			Lower	Upper
1	.30	.114	.11	.48
2	.53	.129	.32	.74
3	1.21	.136	.99	1.44
4	.27	.096	.11	.42
5	.57	.133	.36	.79
Total	.58	.087	.43	.72

There was a statistically significant correlation ($P < 0.001$) between **overall image quality for all cases** and **breast density**. There was an increase in rating with an increase in breast density. Breast density of 1 scored 0.16 (CI 0.08, 0.24), breast density of 2 0.40 (CI 0.30, 0.51), breast density of 3 0.45 (CI 0.35, 0.56) and a breast density of 4 0.55 (CI 0.39, 0.71).

A statistically significant correlation ($P = 0.012$) was shown between **overall lesion quality for all cases** and **breast density**. Breast densities of 1, 2 and 3 were non-inferior, rated 0.41 (CI 0.06, 0.76), 0.62 (CI 0.41, 0.84) and 0.8 (CI 0.53, 1.07) respectively. Breast density of 4 scored 0.14 (CI -0.09, 0.37).

There was a statistically significant correlation ($P = 0.01$) between **benign lesions 'overall lesion quality'** rating and **breast density**. **Breast densities** of 2 and 3 were non-inferior, rated 0.59 (CI 0.35, 0.84) and 0.80 (CI 0.53, 1.08) respectively. Breast densities 1 and 4 scored 0.31 (CI -0.03, 0.66) and 0.06 (CI -0.12, 0.24).

Benign lesions '**scar**' sub-group '**overall lesion quality**' rating were non-inferior for **breast densities** of 2, 3 and 4 with ratings of 0.53 (CI 0.26, 0.8), 0.87 (CI 0.62, 1.13), 0.47 (CI 0.42, 0.53) respectively. Breast density of 1 scored 0.40 (CI -0.18, 0.97).

4. DISCUSSION

The study explored whether synthetic mammograms reconstructed from the 3D data

set of tomosynthesis are non-inferior to true digital mammograms. The results showed that the overall imaging quality of 2D synthesised mammograms was non-inferior to digital mammography. This was shown for 2D synthesised mammograms without and with post processing metal artefact reduction software. There was an increase in scores (+0.01) for overall imaging quality with the use of the software. It should be noted that only 6 cases required post processing metal artefact reduction software. It could be inferred that there would be a greater effect on the improvement of scores if a greater number of cases required the post processing metal artefact reduction.

Similarly the overall lesion quality of 2D synthesised mammograms was non-inferior to digital mammography. This was again shown for 2D synthesised mammograms without and with post processing metal artefact reduction software with an increase in scores (+0.12).

The use of the post processing metal artefact reduction software had a greater effect on overall lesion quality, benign lesion quality and scar subgroup lesion quality than the overall image quality (Figs. 1a, 1b).

This may be explained with the fact that metal artefacts (metallic clips) are in close proximity to lesions after previous surgery or history of breast intervention (5 of the 6 cases were part of the benign lesion "scar" sub group). Even though 2D synthesised mammograms were non-inferior to

digital mammography without the use of post processing metal artefact reduction software, there was a marked improvement in overall

lesion quality rating with the software in use highlighting the importance of using this software in routine imaging acquisition.



Fig. 1a. SM mediolateral oblique view without metal artefact reduction in the region of a benign surgical scar



Fig. 1b. SM mediolateral oblique view with metal artefact reduction in the region of a benign surgical scar

Breast density had a statistically significant impact on the overall imaging quality. It was found that the higher the breast density the better the 2D synthesised mammograms rating, with all ratings being non-inferior.

On analysis of the correlation between breast density and lesion quality it was found that breast densities of 2 and 3 showed lesions the best with 2D synthesised mammograms when compared to digital mammography.

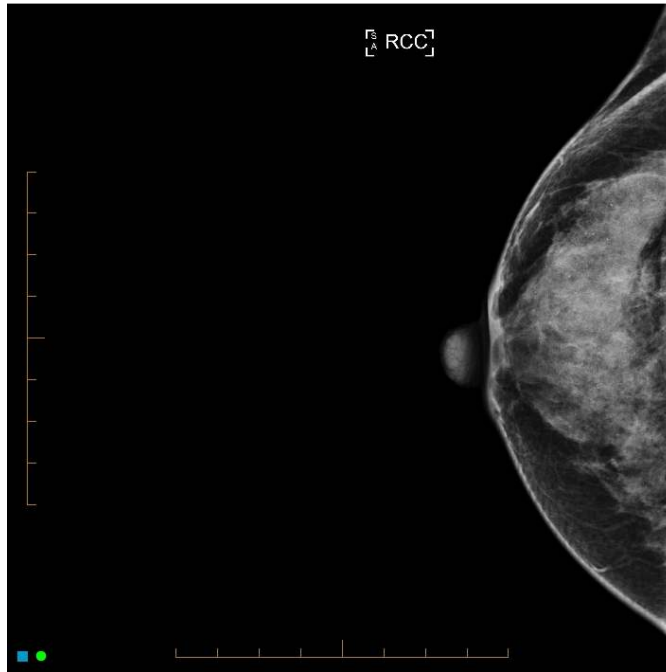


Fig. 2a. FFDM craniocaudal view with widespread benign microcalcifications

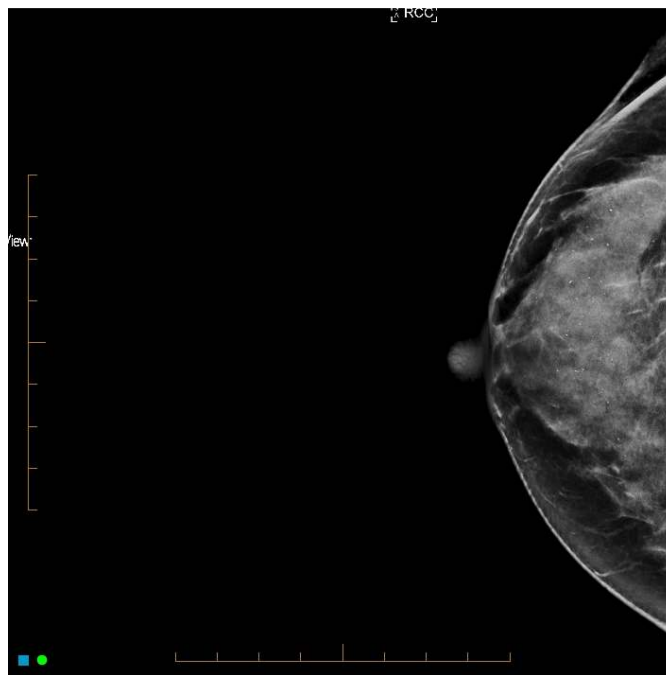


Fig. 2b. SM craniocaudal view with widespread benign microcalcifications

The limitations of the study were that further statistical analysis of lesion subtypes was not possible due to the small number of these cases. This is explained by the fact that the test set of 100 cases was a representative sample of routine reporting in a diagnostic practice. To further statistically investigate subtypes of lesions an enhanced test set would be required with an increased number cases.

The study did not include malignant type microcalcifications. The first reader of the study noted that benign microcalcifications (widespread, round) were better visualised with synthesised 2D mammograms (Figs. 2a, 2b).

In 11 cases benign microcalcifications were identified. All subsequent readers agreed that in these cases benign microcalcifications were better visualised with synthesised 2D mammograms than FFDM. The study was not designed to assess malignant type of microcalcifications and further investigation would be required to determine whether the observation for benign microcalcifications can be extrapolated to malignant type microcalcifications.

Prior to the study, the standard procedure was to acquire FFDM and DBT (synthesised 2D mammograms were available at the radiologist's discretion) in the practise. The results of this study and others [13,14] are encouraging that digital mammography can be eliminated and replaced by 2D synthesised mammograms, resulting in a reduction in radiation dose to that needed for a FFDM.

5. CONCLUSION

The study shows noninferiority of 2D synthesised mammograms when compared to digital mammograms in quality. Synthesised mammograms can replace 2D mammograms when acquired out of a 3D tomosynthesis data set. This allows for both 3D tomosynthesis and 2D mammography interpretation with reduction of the radiation dose by half and reduced time under compression.

Metal artefacts were shown to impact the imaging quality of synthesised 2D mammograms. Routine use of metal artefact reduction software is recommended.

CONSENT

It is not applicable.

ETHICAL APPROVAL

It is not applicable.

ACKNOWLEDGEMENTS

We would like to thank G Coskun (Regional Imaging Tasmania) and S Khrait (Hologic) for IT support.

COMPETING INTERESTS

Authors have declared that no competing interests exist.

REFERENCES

1. Skaane P, Bandos AI, Gullien R, Eben EB, Ekseth U, Haakenaasen U, et al. Comparison of Digital mammography alone and digital mammography plus tomosynthesis in a population-based screening program. *Radiology*. 2013; 267(1):47-56.
2. Baker JA, Lo JY. Breast Tomosynthesis. *Academic Radiology*. 2011;18(10):1298-310.
3. Ciatto S, Houssami N, Bernardi D, Caumo F, Pellegrini M, Brunelli S, et al. Integration of 3D digital mammography with tomosynthesis for population breast-cancer screening (STORM): A prospective comparison study. *Lancet Oncology*. 2013; 14:583-9.
4. Vecchio S, Albanese A, Vignoli P, Taibi A. A novel approach to digital breast tomosynthesis for simultaneous acquisition of 2D and 3D images. *European Radiology*. 2011;21(6):1207.
5. Houssami N, Zackrisson S. Digital breast tomosynthesis: The future of mammography screening or much ado about nothing? *Expert Review of Medical Devices*. 2013;10(5):583.
6. Shah N, Tsz WN, Shaw de Paredes E. Basics of digital breast tomosynthesis. *Applied Radiology*. 2014;43(3):17-204.
7. Skaane P, Bandos AI, Eben EB, Jepsen IN, Krager M, Haakenaasen U, et al. Two-view digital breast tomosynthesis screening with synthetically reconstructed projection images: Comparison with digital breast tomosynthesis with full-field digital mammographic images. *Radiology*. 2014; 271(3):655-63.

8. Hakim CM, Chough DM, Ganott MA, Sumkin JH, Zuley ML, Gur D. Digital breast tomosynthesis in the diagnostic environment: A subjective side-by-side review. *AJR American Journal of Roentgenology*. 2010;195(2):W172-6.
9. Lang K, Andersson I, Zackrisson S. Breast cancer detection in digital breast tomosynthesis and digital mammography-a side-by-side review of discrepant cases. *British Journal of Radiology*. 2014; 87(1040):1.
10. Tagliafico A, Astengo D, Cavagnetto F, Rosasco R, Rescinito G, Monetti F, et al. One-to-one comparison between digital spot compression view and digital breast tomosynthesis. *European Radiology*. 2012; 22(3):539-446.
11. Gur D, Zuley ML, Anello MI, Rathfon GY, Chough DM, Ganott MA, et al. Dose reduction in digital breast tomosynthesis (DBT) screening using synthetically reconstructed projection images: An observer performance study. *Academic Radiology*. 2012;19(2):166-71.
12. Uematsu T. The emerging role of breast tomosynthesis. *Breast Cancer (13406868)*. 2013;20(3):204.
13. Zuley ML, Guo B, Catullo VJ, Chough DM, Kelly AE, Lu AH, et al. Comparison of two-dimensional synthesized mammograms versus original digital mammograms alone and in combination with tomosynthesis images. *Radiology*. 2014;271(3):664-71.
14. Zuckerman SP, Conant EF, Keller BM, Maidment AD, Barufaldi B, Weinstein SP, et al. Implementation of synthesized two-dimensional mammography in a population-based digital breast tomosynthesis screening program. *Radiology*. 2016;160366.
15. Sickles EA, D'Orsi CJ, Bassett LW. *ACR BI-RADS® Atlas, Breast imaging reporting and data system*. Reston, VA: American Journal of Radiology; 2013.

© 2016 Peters et al.; This is an Open Access article distributed under the terms of the Creative Commons Attribution License (<http://creativecommons.org/licenses/by/4.0>), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

Peer-review history:

*The peer review history for this paper can be accessed here:
<http://sciedomain.org/review-history/16893>*