The comparison between synthetic mammography reconstructed from digital breast tomosynthesis and full-field digital mammograms in the detection and assessment of Calcification

C Songsaeng¹, P Woodtichartpreecha¹, P Rohitopakarn¹

¹Department of Radiology, Faculty of Medicine, Prince of Songkla University, Songkla, Thailand
Email: chatsuda.s@psu.ac.th

Abstract. To compare the observer agreement of calcification detection and assessment on synthetic mammography reconstructed to full-field digital mammograms. This retrospective study included 68 patients who undertaken screening mammogram. For each patient, FFDM and DBT were performed at Songklanakarin Hospital. Two-experienced radiologists retrospectively reviewed SM and FFDM, in separate sessions, to detect micro-calcifications and provide a BIRADS assessment. The kappa of SM and FFDM in radiologist1 in type of calcification was 0.49 (0.29-0.68), type of benign was 0.49 (0.29-0.49), type of suspicious was 0.17 (-0.06-0.42). The kappa of SM and FFDM in radiologist2 in type of calcification was 0.74 (0.61-0.87), type of benign was 0.39 (0.2-0.57), type of suspicious was 0.26 (0.11-0.4). The kappa of radiologist1 and radiologist2 of FFDM in type of calcification was 0.49 (0.29-0.68), type of benign was 0.39 (0.2-0.58), type of suspicious was 0.42 (0.24-0.59). The kappa of radiologist1 and radiologist2 of SM in type of calcification was 0.82 (0.72-0.93), type of benign was 0.49 (0.32-0.66), type of suspicious was 0.14 (0.02-0.31). In conclusion, the SM performances are similar as FFDM for detection type of calcification in diagnostic mammography screening whether the radiologist should concern about the image contrast and image reconstruction of SM.

1. Introduction

Since the US Food and Drug Administration (FDA) approved the use of DBT in breast cancer screening in 2011, DBT has been used in combination with full-field digital mammography (FFDM) because of the better accuracy of the combined imaging than FFDM alone. The improved accuracy from DBT plus FFDM may be due to reduced recall rates with simultaneous increase in invasive cancer detection rates as lesion conspicuity has increased. However, the combination of DBT and FFDM (combo) has limitations such increased double radiation dose 2.65 mGy (DBT = 1.45 mGy, FFDM = 1.2 mGy) compared to DBT alone (1.45 mGy).

Recent technologies proposed to resolve this problem use reconstruction algorithms capable of directly generating synthetic mammography reconstruction (SM) 2D image from the DBT dataset, thus avoiding the repeated exposure. These recent studies rate the diagnostic performance of SM supported by the outcome of DBT, as not inferior to FFDM: focusing on the screening environment they reported no significant differences by using 2D synthetic mammography combined with DBT instead of FFDM plus DBT. As we use screening mammography in detection of early breast cancer which usually present with microcalcifications or small masses, for calcifications the morphology as well as distribution are the most important factor in the different between benign and malignant. If synthetic mammography reconstruction (SM) 2D image is equivalent to FFDM in detection of microcalcifications, so we can reduce radiation dose as moving
forward directly to spot magnification view which useful in helping classification of calcifications morphology.

Our study aimed to compare the detection of calcification between the synthetic mammography reconstruction of DBT (SM) and full-field digital mammography (FFDM).

2. Material and methods

2.1 Case selection

The 68 patients who undertaken mammography examination 4 views of both breasts. The number of cases included 136 cases. For each patient, FFDM and DBT were performed between January and June 2016 at Prince of Songklanagarind Hospital. Women aged from 40 years old under general mammogram. Case selection based on FFDM (2D) and BIRADs assessment (selected BIRADS 2–BIRADS 5). All patients had standard 2 views of FFDM and DBT (cranio-caudal and mediolateral oblique views) during single breast compression per view (COMBO-MODE) on digital mammography unit. For each case, synthetic 2D images (in both standard views) were generated from the DBT dataset by using C-view (SM image processing software of Hologic version 1.8.3).

2.2 Methods

Two dedicated breast radiologists who currently interpret a large volume of mammograms in their clinical practice participated in the study. Both were trained in SM reading and had been using SM in recent practice. They both have 6 years-experience to interpret the mammogram in their clinical. The two radiologists reviewed case by separated the side of breast of each patient (right and left) in both standard views (cranio-caudal and mediolateral oblique views). They reviewed and interpreted all cases and blinded clinical information. The SM images blinded to the FFDM interpretation. For each reading and each image modality, the following data were recorded at interpretation, side of breast, type of calcification, type of benign, type of suspicious, view, distribution, position and BIRADS assessment.

2.3 Statistical analysis

The kappa test was used to assess intra-observer and inter-observer variability in terms of final assessment (positive or negative) and visibility scoring for SM and FFDM. Degrees of agreement were categorized as follows: kappa-values of less than 0 was considered to poor agreement; 0.00-0.20 slight agreement; 0.21-0.40, fair agreement; 0.41-0.60, moderate agreement; 0.61-0.80, substantial; and 0.81-1.00, almost perfect agreement.

3. Results

| Morphology          | Radiologist1 | Radiologist2 |
|---------------------|--------------|--------------|
|                     | Kappa  | 95%CI      | Kappa  | 95%CI      |
| Type of calcification | 0.49   | 0.29-0.68  | 0.74   | 0.61-0.87  |
| Type of benign      | 0.49   | 0.29-0.49  | 0.39   | 0.2-0.57   |
| Type of suspicious  | 0.17   | -0.06-0.42 | 0.26   | 0.11-0.4   |
| Type of position    | 0.59   | 0.45-0.74  | 0.64   | 0.48-0.8   |
| Distribution        | 0.75   | 0.75-0.75  | 0.51   | 0.09-0.92  |
| BIRADS              | 0.5    | 0.38-0.62  | 0.6    | 0.51-0.68  |
Table 2. The agreement of inter-observer assessment and visibility scoring of radiologist1 and radiologist2 in SM and FFDM

| Morphology          | SM Kappa | 95%CI       | FFDM Kappa | 95%CI       |
|---------------------|----------|-------------|------------|-------------|
| Type of calcification | 0.82     | 0.72-0.93   | 0.49       | 0.29-0.68   |
| Type of benign      | 0.49     | 0.32-0.66   | 0.39       | 0.2-0.58    |
| Type of suspicious  | 0.14     | -0.02-0.31  | 0.42       | 0.24-0.59   |
| Type of position    | 0.73     | 0.62-0.83   | 0.66       | 0.5-0.83    |
| Distribution        | 0.47     | -0.06-0.99  | 0.46       | -0.09-1.0   |
| BIRADS              | 0.64     | 0.58-0.69   | 0.59       | 0.5-0.69    |

For the detection of calcification, there was moderate agreement between SM and FFDM for radiologist1 and substantial agreement in radiologist2 respectively. For radiologist 1, the kappa was 0.49 (95%CI; 0.29-0.68, \( p < 0.001 \)) indicating a moderate agreement. For radiologist 2, the kappa was 0.74 (95%CI: 0.61-0.87, \( p < 0.001 \)) indicating a substantial agreement. For the detection of benign, there was moderate agreement between SM and FFDM for radiologist1 and fair agreement in radiologist 2 respectively. For radiologist 1, the kappa was 0.49 (95%CI; 0.29-0.49 \( p < 0.001 \)) indicating a moderate agreement. For radiologist 2, the kappa was 0.39 (95%CI: 0.2-0.57, \( p < 0.001 \)) indicating a fair agreement. For the detection of suspicious, there was slight agreement between SM and FFDM for radiologist1 and fair agreement in radiologist 2 respectively. For radiologist 1, the kappa was 0.17 (95%CI; -0.26-0.42, \( p < 0.001 \)) indicating a slight agreement. For radiologist 2, the kappa was 0.26 (95%CI: 0.11-0.4, \( p < 0.001 \)) indicating a moderate agreement. For the detection of BIRADS, there was moderate agreement between the synthetic mammography reconstruction (SM) and full-field digital mammography (FFDM). For radiologist 1, the kappa was 0.5 (95%CI; 0.38-0.62, \( p < 0.001 \)) indicating a moderate agreement. For radiologist 2, the kappa was 0.6 (95%CI: 0.51-0.68, \( p < 0.001 \)) indicating a moderate agreement.

The inter-reader agreement for detection of calcification was almost perfect for using SM and moderate for using FFDM with kappa of 0.82 (95%CI; 0.72-0.93, \( p < 0.001 \)) and 0.49 (95%CI; 0.29-0.68, \( p < 0.001 \)) respectively. The inter-reader agreement for detection of benign was moderate for using SM and fair for using FFDM with kappa of 0.49 (95%CI; 0.32-0.66, \( p < 0.001 \)) and 0.39 (95%CI; 0.2-0.58, \( p < 0.001 \)) respectively. The inter-reader agreement for detection of suspicious was slight for using SM and moderate for using FFDM with kappa of 0.14 (95%CI; -0.02-0.31, \( p < 0.001 \)) and 0.42 (95%CI; 0.24-0.59, \( p < 0.001 \)) respectively.

Figure 1. The kappa between SM and FFDM in Radiologist1 and Radiologist2.
4. Discussion

From table 1, the agreement of intra-observer assessment of SM and FFDM in type of calcification was moderate in radiologist 1 and substantial in radiologist 2, respectively. For the detection of BIRADS, there was moderate agreement between SM and FFDM. The performance of SM and FFDM was found a similar finding of no statistically significant difference in detection of calcification and BIRADS assessment.

Prior studies evaluated the performance of FFDM compared to SM in the screening setting for early T1-stage invasive breast cancer [1,2,3] and determined that SM had similar diagnostic quality to FFDM when used in combination with DBT. Zuley et al. demonstrated that readers were able to identify abnormal microcalcifications 88% of the time on FFDM and 86% of the time on synthetic 2D imaging (p=0.77) [1]. Our study supports their findings with regards to microcalcification detection and BIRADS assessment on SM versus FFDM. From table 2, the agreement of inter-reader assessment of radiologist 1 and radiologist 2 in type of calcification was almost perfect in SM and moderate in FFDM. Due to almost perfect inter-reader agreement, both radiologists had experienced with greater than 6 years of reviewed the cases that were almost perfect agreement in both radiologists.

The limitation of our study was assessed the performance by using SM alone, without DBT information, because we attempted to ensure the validity of the replacement of FFDM with SM in DBT based screening by evaluating the independent performance of SM in comparison with FFDM. However, as 2D SM is derived from DBT images, we think that 2D SM will not be interpreted alone, but along with DBT images in clinical practice. Our results show that the SM showed the similar performance to standard FFDM for detection type of calcification whether we should concern about the image contrast setting and image reconstruction of SM.

5. Conclusion

The SM performances are similar as FFDM for detection type of calcification in diagnostic mammography screening whether the radiologist should concern about the image contrast and image reconstruction of SM.
6. References

[1] Zuley ML, Guo B and Catullo VJ 2014 Radiology 271 664
[2] Skaane P, Bandos AI and Eben EB 2014 Radiology 271 655
[3] Zuckerman SP, Maidment AD and Weinstein SP 2017 AJR 209 222
[4] Choi J, Han B and Young Ko E 2016 Eur. Radiol. 26 2538
[5] Giovanna M, Manuela D and Nehmat H 2017 Breast Cancer Res Treat. 8 4458
[6] Margarita L, Ben G and Victor J 2014 Radiology. 271 664
[7] Spangler ML, Zuley ML and Sumkin JH 2011 AJR 196 743

Acknowledgments

The authors wish to thank Assoc. Prof. Chanon Kongkamol, Prince of Songkla University, Songkhla, Thailand for his suggestions and statistics analysis.