ORIGINAL ARTICLE

Radiation Medicine: Vol. 22 No. 6, 408–412 p.p., 2004

# Comparison of Screen-film and Full-field Digital Mammography in Japanese Population-based Screening

**Takayuki Yamada,* Mioko Saito,* Tadashi Ishibashi,* Masahiro Tsuboi,* Toshio Matsuhashi,* Akihiro Sato,* Haruo Saito,* Shoki Takahashi,* Koji Onuki,** and Noriaki Ouchi****

***Purpose:*** To investigate how the greater contrast of full-field digital mammography (FFDM) affects the detection of suspicious lesions in Japanese population-based screening.
***Materials and Methods:*** Screen-film mammography (SFM) and FFDM were performed in 480 women aged 50 years or more. A set of mediolateral oblique views was obtained with each modality. All mammograms were independently double-read. The five-scale category assessment and type of finding using the Breast Imaging Reporting and Data system (BI-RADS) nomenclature were given. Intraobserver variance, recall rates, and positive predictive value were calculated.
***Results:*** The findings between the two modalities were discordant. $\kappa$-values for each reader were 0.619 and 0.385, respectively. Almost half of the microcalcifications were called with both modalities. The detection of masses was less concordant between the readers (27%). The masses were detected more frequently with FFDM (73%). Other findings were only detected with one modality. The recall rate was not significantly different (2.9% with SFM vs. 4.2% with FFDM; $p=0.253$). The positive predictive value was not significantly different (14% with SFM vs. 10% with FFDM; $p=0.69$), either. Two patients with breast cancer were detected with both modalities.
***Conclusion:*** Recall rates and positive predictive value were not significantly different between SFM and FFDM. Cancers were detected with both modalities.

***Key words:*** screening mammography, screen-film mammography (SFM), full-field digital mammography (FFDM), breast cancer

## INTRODUCTION

TWO STUDIES EXAMINED THE MAJORITY OF RANDOMIZED controlled studies for screening mammography and concluded that screening mammography is ineffective in reducing breast cancer mortality.[1,2] However, the investigators defended their trials, and criticisms of all but one of the trials excluded from the meta-analysis have been answered.[3] Screening mammography will definitely continue to play an important role in detecting early breast cancer. In Japan, screening mammography began to be combined with the conventional checkup for breast cancer by physician's visual inspection and palpation from the year 2000. Women aged 50 years and more are recommended for medical checkup with screening mammography.

Screen-film mammography (SFM) has been used to detect suspicious lesions for screening. Although SFM is widely used, it has important limitations in detecting subtle soft-tissue lesions, especially in the presence of dense glandular tissue. Currently, full-field digital mammography (FFDM) is commercially available. FFDM is expected to identify at least some cancers in dense lesions since the contrast of FFDM is superior to SFM.[4] However, discordant reading between them was frequent in the studies comparing SFM and prototype FFDM.[5,6] The most recent study using a commercially available indirect detector system reported that there was no significant difference between the two modalities.[7]

Received November 21, 2003; revision accepted May 17, 2004.

Departments of *Diagnostic Radiology and **Surgical Oncology, Tohoku University, Graduate School of Medicine

Reprint requests to Takayuki Yamada, Department of Diagnostic Radiology, Tohoku University, Graduate School of Medicine, 1-1 Seiryomachi, Aoba-ku, Sendai, Miyagi 980-8574, JAPAN.

Because a set of mediolateral oblique (MLO) views is only obtained in the Japanese screening setting, it remains unresolved as to how greater contrast affects the detection of suspicious lesions in screening mammography. These modalities have not been compared previously in Japanese screening.

In this study, we performed both SFM and FFDM for the same Japanese population in medical checkups and evaluated whether FFDM is superior to SFM.

## Materials and Methods

### *Patient selection and data acquisition*

This study was approved by the Institutional Review Board of our university. Local women aged 50 years and more who were assigned to our institute for screening mammography were invited. Although a local organization has had a voluntary breast cancer screening service, this was the first population-based screening program for local people. Before screening examination, the participants were informed about this study. Since all participants gathered at the same time, technologists took the FFDM pictures first. Immediately after that, the same technologist took the SFM pictures in those who consented to the additional examinations by written form. This study was performed from July 2001 to November 2001. Since the original hard-copies of FFDM were submitted to the screening center, the images of FFDM were retrieved and printed on hard-copies again for this research. SFM images were stored in our institute. No cancer patients were found at the time of the review in September 2002. Two patients with breast cancer that underwent both SFM and FFDM were added randomly by one of the authors. Both cancers were diagnostic cases and not screen-detected. Finally, we evaluated 480 people. No cancer patients were found from the group of attendees as of September 2003.

### *Equipment*

FFDM was performed with a commercially available unit Senographe 2000D (GE Medical Systems, Milwaukee, WI). Technical factors (peak kilovoltage, milliampere-second, target, and filter) were automatically determined by pre-exposure of the breasts (automated optimization parameters). The standard mode in which all parameters were set for the routine examination was used for all cases. The combinations of target and filter included rhodium/rhodium, molybdenum/rhodium, and molybdenum/molybdenum. SFM was performed with a Mammomat 3000 Nova (Siemens Medical Systems, Solna, Sweden). The unit has suggested technical parameters, but technicians can change them manually. The combinations of target and filter include molybdenum/molybdenum, molybdenum/rhodium, and tungsten/molybdenum. Therefore, the parameters were not identical for SFM and FFDM. All SFM and FFDM images were printed on hard-copies. FFDM images were printed by a commercially available laser imager unit (Kodak DryView 8610; Eastman Kodak, Rochester, NY). Pixel size of the printer was 39 $\mu$m, and the size of the film was 8×11 inch. Processed images were obtained using the manufacturer's processing algorithm. The window was set with the initial set-up parameters with the level at about 2300 and width at 750. A commercial screen-film system (Kodak Min-R 2000; Eastman Kodak, Rochester, NY) was used for SFM. The quality control procedures used for SFM were in accordance with the Japanese Mammography Quality Control Manual.

### *Mammogram interpretation*

A blinded review was performed independently by two readers. Both readers were radiologists from the body imaging section of the same institute. The experience of both readers regarding reading mammograms was similar over a period of five years. They were qualified for reading screening mammograms in Japan. Before they reviewed the mammograms, they were told that cases of breast cancer had been added but they did not know the number of cases. The readers read the mammograms in a darkened room on a standard mammography light-box. They read SFM first, and after at least one month they read FFDM. They gave the following information:

(1) A description of the type and location of the findings obtained using the Breast Imaging Reporting and Data system (BI-RADS) nomenclature.

(2) Five-scale category assessment from 1 to 5. Category assessment except 3 is the same as BI-RADS. BI-RADS 0 has not been adopted in the Japanese screening setting, and such cases are placed in category 3. Thus, the findings in category 3 require additional work-up that is different from the original BI-RADS 3, which recommends short-term follow-up after thorough additional work-up.

### *Statistical analysis*

The significance of the recall rate was analyzed using the McNemar chi-square test. The significance of the positive predictive value was evaluated using Fisher's exact test. $\kappa$-value was calculated with regard to the intra-observer variance between SFM and FFDM for each reader.

**Table 1. Distribution of findings detected by each modality**

| | | Both | SFM only | FFDM only | Total |
|---|---|---|---|---|---|
| Reader 1 | Mass | 6 (35) | 4 (24) | 7 (41) | 17 (100) |
| | Calcification | 9 (69) | 1 (8) | 3 (23) | 13 (100) |
| | Others | 0 (0) | 0 (0) | 2 (100) | 2 (100) |
| | Subtotal | 15 (47) | 5 (16) | 12 (37) | 32 (100) |
| Reader 2 | Mass | 3 (19) | 5 (31) | 8 (50) | 16 (100) |
| | Calcification | 10 (42) | 8 (33) | 6 (25) | 24 (100) |
| | Others | 0 (0) | 3 (43) | 4 (57) | 7 (100) |
| | Subtotal | 13 (28) | 16 (34) | 18 (38) | 47 (100) |
| | Total | 28 (35) | 21 (27) | 30 (38) | 79 (100) |

Numbers in parentheses are percentages.
$\kappa$-values for reader 1 and reader 2: 0.619 and 0.385, respectively.
SFM = screen-film mammography
FFDM = full-field digital mammography

**Table 2. Recall rate and distribution of category assessment by each modality**

| | SFM | | FFDM | | Total |
|---|---|---|---|---|---|
| | Reader 1 | Reader 2 | Reader 1 | Reader 2 | |
| Category 3 | 9 | 12 | 11 | 16 | 48 |
| Category 4 | 2 (2) | 3 (1) | 5 (1) | 3 | 13 (4) |
| Category 5 | 0 | 2 (1) | 2 (1) | 3 (2) | 7 (4) |
| Total | 11 | 17 | 18 | 22 | 68 |
| Recall rate* | 2.90% | | 4.20% | | |

*: McNemar $\chi^2$ test, p=0.253
Numbers in parentheses are cancer patients.
SFM = screen-film mammography
FFDM = full-field digital mammography

## RESULTS

The mean dose was 1.48 mGy for each exposure with FFDM. This value was obtained from all participants who underwent FFDM for screening. Although we did not compare the dose with the two modalities, the last study we did for diagnostic mammography[10] showed that the mean peak kilovoltage and milliampere-second on SFM and FFDM are close to each other.

The distribution of findings detected with each modality was quite different between the two readers. Nearly half of the findings were detected with both SFM and FFDM by the first reader (15 of 32 findings, 47%; Table 1). The percentage of findings detected with only SFM was small (5 of 32 findings, 16%; Table 1). $\kappa$-value was 0.619. On the other hand, the findings of each group were more evenly distributed by the second reader (both, only SFM, and only FFDM: 28%, 34%, and 38%, respectively; Table 1). $\kappa$-value was 0.385.

Nineteen of the 37 findings (51%) of microcalcifications were called with both modalities; however, the other 18 findings were still called with only one modality. Nine (24%) of the 18 findings were detected with each modality (Table 1). The detection of masses was less concordant between the two modalities as compared with that of microcalcifications (9 of 33 findings, 27%; Table 1). The masses were detected more frequently with FFDM (24 of 33 findings, 73%; Table 1). Other findings including architectural distortion and asymmetric density were not concordant with the modalities (Table 1).

The mean recall rate with FFDM was slightly higher than that with SFM (SFM vs. FFDM: 2.9% vs. 4.2%; Table 2). One reader doubled the recall rate from 1.9% to 3.8%. However, significant difference was not demonstrated (p=0.253, McNemar $\chi^2$). Of the recalled cases, the majority were category 3 (48 of 68 cases, 70%; Table 2), and they were considered to be necessary for

additional imaging work-up. The numbers of category 4 and 5 cases were almost the same for the first reader (28% with SFM vs. 29% with FFDM). On the other hand, the rates for these categories definitely increased with the second reader (18% with SFM vs. 39% with FFDM). The positive predictive value of screening is defined as the fraction of recalled examinations that lead to a diagnosis of breast cancer. The positive predictive value with FFDM decreased as compared with SFM (14% with SFM vs. 10% with FFDM; Table 3). However, no significant difference was noted (p=0.69, Fisher's exact test; Table 3). The cases of breast cancer included in this study were detected with both modalities by both readers. One case was ductal carcinoma *in situ* and the other was invasive ductal carcinoma. A patient rated as category 5 except on SFM by the first reader underwent biopsy in the actual diagnostic work-up. The pathologic specimen showed no malignant cells and secretory microcalcifications were seen in the duct in part of the specimen.

**Table 3. Positive predictive value with each modality**

| | Cancer | | Total |
|---|---|---|---|
| | (+) | (−) | |
| SFM positive | 4 (14) | 24 | 28 (100) |
| FFDM positive | 4 (10) | 36 | 40 (100) |
| Total | 8 | 60 | 68 |

Numbers in parentheses are percentages.
Fisher's exact test, p=0.69.
SFM = screen-film mammography
SFM positive = cases of category 3, 4, and 5 with SFM
FFDM = full-field digital mammography
FFDM positive = cases of category 3, 4, and 5 with FFDM

## Discussion

In this study, FFDM was used for screening mammography of a local population. This is the first trial comparing SFM and FFDM in a Japanese screening setting.

The findings were discordant between the two modalities with each reader. Although about half of the findings were concordant for one reader, the concordance rate decreased into 30% for the other reader. In prior studies using prototype FFDM, discordant readings between SFM and FFDM were not uncommon in screening mammography.[5,6] In a past report, findings detected with both modalities only represented 13.5%, and half of the findings were only detected with SFM.[5] Another study by the same author reported the most common reason for findings detected only with FFDM was fortuitous positioning.[6] In a diagnostic setting, the difference in lesion visibility was a minor reason,[8] but this is probably because the women had clinical concerns such as clinical symptoms and abnormal findings. Findings seen on only one standard projection at screening mammography are reported to be 3.3%, and about half of them were able to be correctly assessed as representing superimposition of normal structure.[9] If different positioning occurred in our cases, only one set of MLO would make it difficult to judge superimposition of normal tissue. This also could increase discordant readings.

Findings of masses were picked up more frequently by FFDM than by SFM. The greater contrast of FFDM may be a cause. The difference in optical density is actually larger on FFDM than on SFM in a diagnostic setting.[10] In contrast, microcalcifications were picked up by both modalities more frequently than masses. This is probably because microcalcifications seem easier to recognize than masses or asymmetric density. In a diagnostic setting, the visibility of microcalcifications on FFDM was equivalent or superior to SFM.[4,10] In addition, one prior report noted that FFDM with a 100-$\mu$m pixel size provides better image quality than SFM in more than 50% of cases.[11]

FFDM slightly increased the recall rate, from 2.9% to 4.2%. One reader doubled the recall rate from 1.9% to 3.8%. However, the difference was not significant. In prior studies, the recall rate with FFDM did not show this unique trend. In studies by the same author, the recall rate was decreased.[5,6] On the other hand, other study showed a slight increase from 3.5% to 4.6%.[7] Therefore, recall rates with FFDM do not necessarily show a unique trend, and factors influencing the recall rates cannot be simply ascribed to difference in modalities. One limitation of this study was that it was a retrospective experimental study. This means that our reading results with regard to the recall rate did not affect the actual decision for further work-up. Thus, we cannot conclude whether or not FFDM actually increased unnecessary biopsy. To determine this, it would be necessary to investigate the reading results at the screening center.

Both the cancers were depicted on both modalities; however, it would be impossible to conclude that FFDM would not be useful to detect more cases of breast cancer in a Japanese screening setting. This is because no cancer patients were found in this group of people; only two cases of breast cancer were added considering the prevalence of breast cancer in the previous study performed for the volunteers in this area.[12] In addition, the cases of breast cancer were diagnostic cases, not screen-detected cancers. This means that findings were definite. This is a limitation of this study. A prior report from Norway described that there was no statistically

significant difference in cancer detection rate between SFM and FFDM.[7] The American College of Radiology Imaging Network (ACRIN) Digital Mammographic Imaging Screening Trial (DMIST) will enroll 49,500 women under a protocol that requires both digital and conventional SFM at more than 20 participating centers in North America.[13] The ACRIN study is expected to offer evidence concerning the effectiveness of FFDM for screening mammography. However, Japanese screening mammography is unique in that one set of MLO is obtained. Therefore, similar study might be necessary to evaluate the effectiveness of FFDM in Japanese population-based screening mammography.

There are other shortcomings in this study. The first is that we did not record the density of the breast tissue since we only recorded the concerned finding in accordance with the form at the screening center. It merely required a check of abnormal findings, and did not require the density of the breast tissue. This made it impossible to evaluate how dense breast affects the detection of breast cancer on FFDM. The second is that we did not record the age of the participants in this study. The individual data of the participants, except their names, were not left in our institute, and this prevented us from knowing the exact distribution of participants' ages. However, we believe the distribution of ages in this study was not significantly different from the population who underwent the screening.

In conclusion, findings with SFM and FFDM were discordant. Nearly half of the microcalcifications were detected with both modalities, but masses were more frequently detected with FFDM. Recall rates and positive predictive value were not significantly different between them. Cancers were depicted on both modalities; however, a larger study will be necessary to evaluate the effectiveness of FFDM in Japanese screening mammography.

## Acknowledgements

We acknowledge the important contributions of the radiologic technologists, especially Masahiro Sai.

## References

1) Gotzsche PC, Olsen O. Is screening for breast cancer with mammography justifiable? *Lancet,* 355: 129–134, 2000.
2) Olsen O, Gotzsche PC. Cochrane review on screening for breast cancer with mammography. *Lancet,* 358: 1340–1342, 2001.
3) Fletcher SW, Elmore JG. Clinical practice. Mammographic screening for breast cancer. *N Engl J Med,* 348: 1672–1680, 2003.
4) Obenauer S, Luftner-Nagel S, von Heyden D, Munzel U, Baum F, Grabbe E. Screen film vs full-field digital mammography: image quality, detectability and characterization of lesions. *Eur Radiol,* 12: 1697–1702, 2002.
5) Lewin JM, Hendrick RE, D'orsi CJ, *et al.* Comparison of full-field digital mammography with screen-film mammography for cancer detection: results of 4945 paired examinations. *Radiology,* 218: 873–880, 2001.
6) Lewin JM, D'orsi CJ, Hendrick RE, *et al.* Clinical comparison of full-field digital mammography and screen-film mammography for detection of breast cancer. *AJR Am J Roentgenol,* 179: 671–677, 2002.
7) Skaane P, Young K, Skjennald A. Population-based mammography screening: comparison of screen-film and full-digital mammography with soft-copy reading – Oslo I study. *Radiology,* 229: 877–884, 2003.
8) Venta LA, Hendrick RE, Adler YT, *et al.* Rates and causes of disagreement in interpretation of full-field digital mammography and film-screen mammography in a diagnostic setting. *AJR Am J Roentgenol,* 176: 1241–1248, 2001.
9) Sickles EA. Findings at mammographic screening on only one standard projection: outcomes analysis. *Radiology,* 208: 471–475, 1998.
10) Yamada T, Ishibashi T, Sato A, *et al.* Comparison of screen-film and full-field digital mammography: image contrast and lesion characterization. *Radiat Med,* 21: 166–171, 2003.
11) Fischer U, Baum F, Obenauer S, *et al.* Comparative study in patients with microcalcification: full-field digital mammography vs screen-film mammography. *Eur Radiol,* 12: 2679–2683, 2002.
12) Saito M, Ohnuki K, Yamada T, *et al.* Efficacy of computer-aided detection system for screening mammography. *J Jpn Assoc Breast Cancer Screen*, 11: 251–256, 2002. (in Jpse.)
13) Welsh CR, Young B, Gopalakrishnan V, *et al.* ACRIN digital mammographic imaging screening trial informatics infrastructure. In *Computer Assisted Radiology and Surgery,* (Lemke HU, Vannier MW, Inamura K, Farman AG, Doi K, Reiber JHC eds.; Springer, Germany), pp. 503–508, 2002.