



Effect of Different Types of Mammography Equipment on Screening Outcomes: A Report by the Alliance for Breast Cancer Screening in Korea

Bo Hwa Choi, MD, PhD¹, Eun Hye Lee, MD, PhD², Jae Kwan Jun, MD, PhD³, Keum Won Kim, MD, PhD⁴, Young Mi Park, MD, PhD⁵, Hye-Won Kim, MD, PhD⁶, You Me Kim, MD, PhD⁷, Dong Rock Shin, MD⁸, Hyo Soon Lim, MD, PhD⁹, Jeong Seon Park, MD, PhD¹⁰, Hye Jung Kim, MD¹¹;
Alliance for Breast Cancer Screening in Korea (ABCS-K)

¹Department of Radiology, Gyeongsang National University School of Medicine and Gyeongsang National University Changwon Hospital, Changwon, Korea; ²Department of Radiology, Soonchunhyang University Bucheon Hospital, Soonchunhyang University College of Medicine, Bucheon, Korea; ³National Cancer Control Institute, National Cancer Center, Goyang, Korea; ⁴Department of Radiology, Konyang University Hospital, Konyang University College of Medicine, Daejeon, Korea; ⁵Department of Radiology, Busan Paik Hospital, Inje University College of Medicine, Busan, Korea; ⁶Department of Radiology, Wonkwang University Hospital, Wonkwang University School of Medicine, Iksan, Korea; ⁷Department of Radiology, Dankook University Hospital, Dankook University College of Medicine, Cheonan, Korea; ⁸Department of Radiology, Gangneung Asan Hospital, University of Ulsan College of Medicine, Gangneung, Korea; ⁹Department of Radiology, Chonnam National University Hwasun Hospital, Chonnam National University College of Medicine, Hwasun, Korea; ¹⁰Department of Radiology, Hanyang University Hospital, Hanyang University College of Medicine, Seoul, Korea; ¹¹Department of Radiology, Kyungpook National University Medical Center, Kyungpook National University College of Medicine, Daegu, Korea

Objective: To investigate the effects of different types of mammography equipment on screening outcomes by comparing the performance of film-screen mammography (FSM), computed radiography mammography (CRM), and digital mammography (DM).

Materials and Methods: We retrospectively enrolled 128756 sets of mammograms from 10 hospitals participating in the Alliance for Breast Cancer Screening in Korea between 2005 and 2010. We compared the diagnostic accuracy of the types of mammography equipment by analyzing the area under the receiver operating characteristic curve (AUC) with a 95% confidence interval (CI); performance indicators, including recall rate, cancer detection rate (CDR), positive predictive value₁ (PPV₁), sensitivity, specificity, and interval cancer rate (ICR); and the types of breast cancer pathology.

Results: The AUCs were 0.898 (95% CI, 0.878–0.919) in DM, 0.860 (0.815–0.905) in FSM, and 0.866 (0.828–0.903) in CRM ($p = 0.150$). DM showed better performance than FSM and CRM in terms of the recall rate (14.8 vs. 24.8 and 19.8%), CDR (3.4 vs. 2.2 and 2.1 per 1000 examinations), PPV₁ (2.3 vs. 0.9 and 1.1%), and specificity (85.5 vs. 75.3 and 80.3%) ($p < 0.001$) but not in terms of sensitivity (86.3 vs. 87.4 and 86.3%) and ICR (0.6 vs. 0.4 and 0.4). The proportions of carcinoma *in situ* (CIS) were 27.5%, 13.6%, and 11.8% for DM, CRM, and FSM, respectively ($p = 0.003$).

Conclusion: In comparison to FSM and CRM, DM showed better performance in terms of the recall rate, CDR, PPV₁, and specificity, although the AUCs were similar, and more CISs were detected using DM. The application of DM may help to improve the quality of mammography screenings. However, the overdiagnosis issue of CIS using DM should be evaluated.

Keywords: Breast neoplasms; Digital mammography; Screening; Sensitivity and specificity

Received January 4, 2019; accepted after revision July 29, 2019.

This study was supported by a grant from the National R&D Program for Cancer Control, Ministry of Health & Welfare, Republic of Korea (1520200).

Corresponding author: Eun Hye Lee, MD, PhD, Department of Radiology, Soonchunhyang University Bucheon Hospital, Soonchunhyang University College of Medicine, 170 Jomaru-ro, Bucheon 14584, Korea.

• Tel: (8232) 621-5851 • Fax: (8232) 621-6950 • E-mail: grace@schmc.ac.kr

This is an Open Access article distributed under the terms of the Creative Commons Attribution Non-Commercial License (<https://creativecommons.org/licenses/by-nc/4.0>) which permits unrestricted non-commercial use, distribution, and reproduction in any medium, provided the original work is properly cited.

INTRODUCTION

The incidence of breast cancer in Korea has been increasing, and currently, breast cancer is the most common cancer among Korean women after thyroid cancer (1). The incidence of breast cancer in Korea is the highest among women in their late 40s, which is approximately 10 years less than the typical age of incidence in Western countries. Unfortunately, the proportion of women with dense breast tissue is high in women in their 40s and 50s; therefore, early detection of breast cancer at these ages is difficult (2).

Korean women aged 40 years or older have been advised to undergo biannual mammography screenings under the auspices of the National Cancer Screening Program (NCSP) since 1999. A recent study conducted by the Alliance for Breast Cancer Screening in Korea (ABCS-K) showed that the sensitivity and cancer detection rate (CDR) during mammography screenings were comparable to those obtained during screening programs in Western countries (3). However, the recall rate, positive predictive value (PPV), and specificity were suboptimal compared to those reported in Western studies. These differences could be due to various factors such as examinees, radiologists, and equipment.

Until now, many studies on the impact of screening with digital mammography (DM) have shown that the diagnostic accuracy of DM is similar to that of film-screen mammography (FSM). It is noteworthy that several studies have shown that the use of DM for screening purposes could improve the diagnostic performance significantly compared to FSM among women under the age of 50 and women with dense breasts (4-6). However, some studies have argued that DM has inferior recall rates and PPV (7-10). To our knowledge, no comparative studies on the impact of DM have been conducted among Asian women, who tend to have dense breasts (11-13).

The study period was from 2005 to 2010 during the ABCS-K, which was the transitional time from FSM or computed radiography mammography (CRM) to DM in Korea. Thus, we aimed to investigate how the different types of mammography equipment affect the screening outcomes in Korean women.

MATERIALS AND METHODS

Study Population and Data Collection

The study population and data collection methods were

fully described in our previous report (3). Ten university-affiliated hospitals which participated in both the NCSP and the ABCS-K were involved in this retrospective study. This study was approved by the Institutional Review Boards of all participating hospitals, and the need for informed participant consent was waived.

We collected information about participants, radiologists, and the mammography equipment that was used from all participating hospitals between January 2005 and December 2010. Our research database was matched with the databases of the NCSP and the National Health Insurance Service (NHIS), where the results of mammography screenings and cancer outcomes were available, respectively. We extracted the number of enrolled cases according to the types of mammography equipment used: FSM, CRM, and DM.

We considered the American College of Radiology Breast Imaging-Reporting and Data System (ACR BI-RADS) categories 1 and 2 as negative results and categories 0, 4, and 5 as positive results (14). We obtained information about cancer diagnoses from the NHIS database until December 2011 to account for the 12-month period after the screening. Breast cancers included both invasive cancer and carcinoma *in situ* (CIS).

Outcome Measurements and Statistical Analysis

The diagnostic accuracy in the detection of breast cancer was assessed through calculation of the area under the curve (AUC) and the corresponding 95% confidence intervals (CI) from a receiver operating characteristic (ROC) curve analysis for FSM, CRM, and DM. The AUCs were compared using the method described by DeLong et al. (15).

We also calculated performance indicators, including the recall rate, CDR per 1000 examinations, PPV₁, sensitivity, specificity, false positive rate (FPR), and interval cancer rate (ICR) per 1000 negative examinations and compared them according to the types of equipment used. The performance indicators were defined according to the ACR BI-RADS (14). The recall rate was calculated as the percentage of examinations that were screened and recalled for further evaluation. The CDR was calculated as the number of breast cancer cases detected per 1000 examinations. The interval cancers were defined as histology-proven invasive or *in situ* cancers found within one year following the negative screening. In addition, the ICR was calculated as the number of interval cancer incidents per 1000 negative examinations. The PPV₁ refers to the percentage of all positive screening examinations that resulted in histology-

proven breast cancers within one year following the screening. We used the Wald chi-squared test to compare performance indicators according to the types of equipment used, and we performed logistic regression analysis to control the characteristics of radiologists, including subspecialty and level of experience in breast imaging. We also compared the proportion of invasive and *in situ* cancers according to the types of mammography equipment used.

All *p* values less than 0.05 were considered statistically significant. All statistical analyses were conducted using SAS software, version 9.2 (SAS Institute Inc., Cary, NC, USA).

RESULTS

Patient Characteristics

The baseline raw data of the ABCS-K from the 10 participating hospitals from 2005 to 2010 comprised 130537 sets of mammograms. Among them, 1121 sets of mammograms with incomplete identification numbers, 557 sets of mammograms from women with a previous diagnosis of breast cancer, and 103 sets of mammograms showing interstitial mammoplasty were excluded. Finally, 128756 sets of mammograms from 103411 women were included in the research database for this study. Among all examinations, 62.9% (80958 of 128756) was the first (prevalence) screening and 37.1% (47798 of 128756) was subsequent (incidence) screening.

Four hundred breast cancer patients, including 322 invasive cancers and 78 CIS, were registered in the database of the NHIS within one year following the screening. Among them, 346 (86.5%) patients had screening-detected cancers and 54 (13.5%) patients had interval cancers.

Equipment Characteristics

During the ABCS-K, seven FSM systems, five CRM systems, and seven DM systems were used. The detailed information and the number of mammography equipments used are summarized in Table 1.

Among the 128756 sets of mammograms, 33979 (26.4%) were obtained using FSM, 41697 (32.4%) were obtained using CRM, and 53080 (41.2%) were obtained using DM. The proportion of women who underwent DM increased significantly throughout the study period (Fig. 1). The distribution of the cases according to the institution and the type of mammography equipment used is shown in Figure 2.

Table 1. Mammography Systems Used in Alliance for Breast Cancer Screening in Korea

Type	Model and Manufacturer	n
FSM (n = 7)	Senographe DMR (GE Healthcare)	2
	Senographe 500T (GE Healthcare)	1
	Senographe 600T (GE Healthcare)	1
	Lorad M-IV (Hologic)	1
	MAMMOMAT 300 (Siemens Healthineers)	1
	PERFORMA MHR-35-P (Instrumentarium)	1
CRM (n = 5)	FCR + Senographe DMR (GE Healthcare)	3
	FCR + MAMMOMAT 300 (Siemens Healthineers)	1
	FCR + PERFORMA MHR-35-P (Instrumentarium)	1
DM (n = 7)	Senographe 2000D (GE Healthcare)	2
	Senographe DS (GE Healthcare)	2
	Lorad Selenia (Hologic)	3

CRM = computed radiography mammography, DM = full-field digital mammography, FCR = FCR CAPSULA XL II (Fujifilm), FSM = film-screen mammography, n = number of mammography systems

Diagnostic Performance

DM showed the largest AUC (AUC = 0.898; 95% CI: 0.878, 0.919) compared to FSM (AUC = 0.860; 95% CI: 0.815, 0.905) and CRM (AUC = 0.866; 95% CI: 0.828, 0.903) (Fig. 3). However, there was no significant difference in AUCs between the types of mammography equipment used (*p* = 0.150).

For performance indicators, DM showed better performance than FSM and CRM in terms of recall rate (14.8 vs. 24.8 and 19.8%, respectively), CDR (3.4 vs. 2.2 and 2.1, respectively), PPV₁ (2.3 vs. 0.9 and 1.1%, respectively), specificity (85.5 vs. 75.3 and 80.3%, respectively), and FPR (14.5 vs. 24.7 and 19.7%, respectively) after adjustment for radiologist factors (*p* < 0.001) (Table 2). However, there were no significant differences in sensitivity (86.3 vs. 87.4 and 86.3%, respectively; *p* = 0.819) and ICR (0.6 vs. 0.4 and 0.4; *p* = 0.187) among the types of equipment used.

Types of Cancers

Among the 346 patients with screening-detected cancers, 275 patients had invasive cancers and 71 patients had *in situ* cancers. The numbers of patients with screening-detected cancers and types of pathology according to the types of mammography equipment used are summarized in Table 3. The proportions of invasive cancers and CIS in screening-detected cancers were 88.2% and 11.8%, 86.4% and 13.6%, and 72.5% and 27.5% for FSM, CRM, and DM, respectively. DM detected more CIS than FSM and CRM (*p* = 0.003).

DISCUSSION

In our study, we found that DM qualitatively showed better diagnostic performance than FSM and CRM in terms

of recall rate, CDR, PPV₁, specificity, and FPR, although sensitivity, ICR, and overall diagnostic accuracy were not significantly different among all types of equipment. We also found that DM detected a higher proportion of CIS than

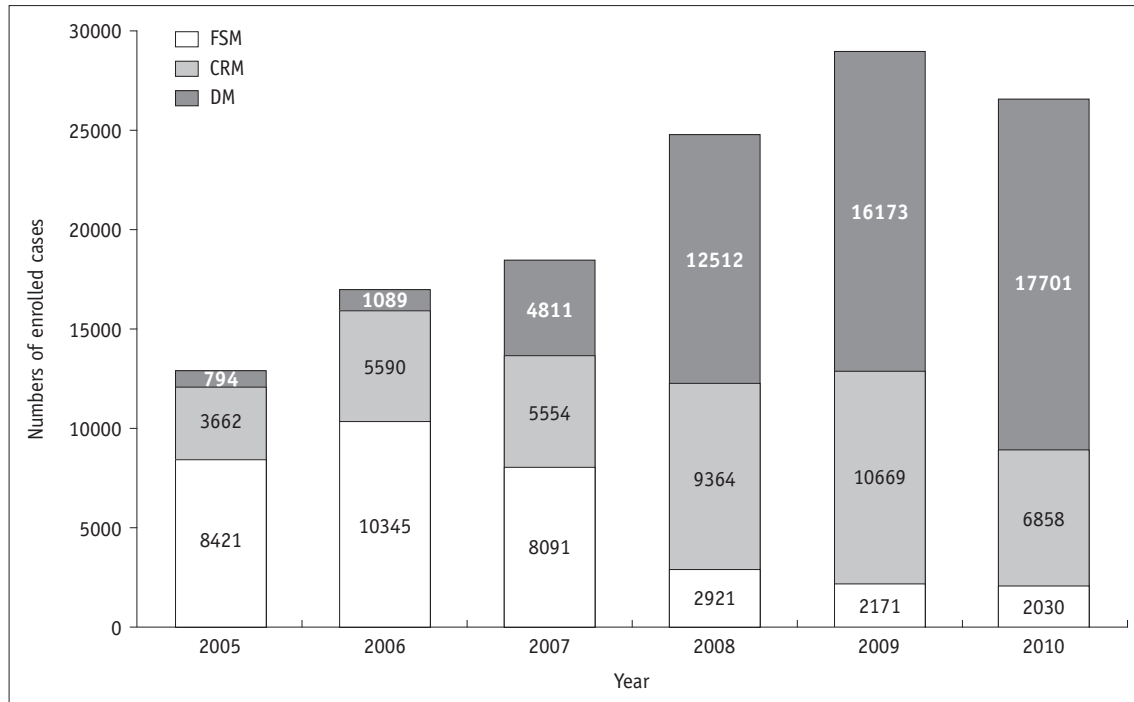


Fig. 1. Distribution of enrolled cases throughout study period by type of mammography equipment used in ABCS-K. ABCS-K = Alliance for Breast Cancer Screening in Korea, CRM = computed radiography mammography, DM = full-field digital mammography, FSM = film-screen mammography

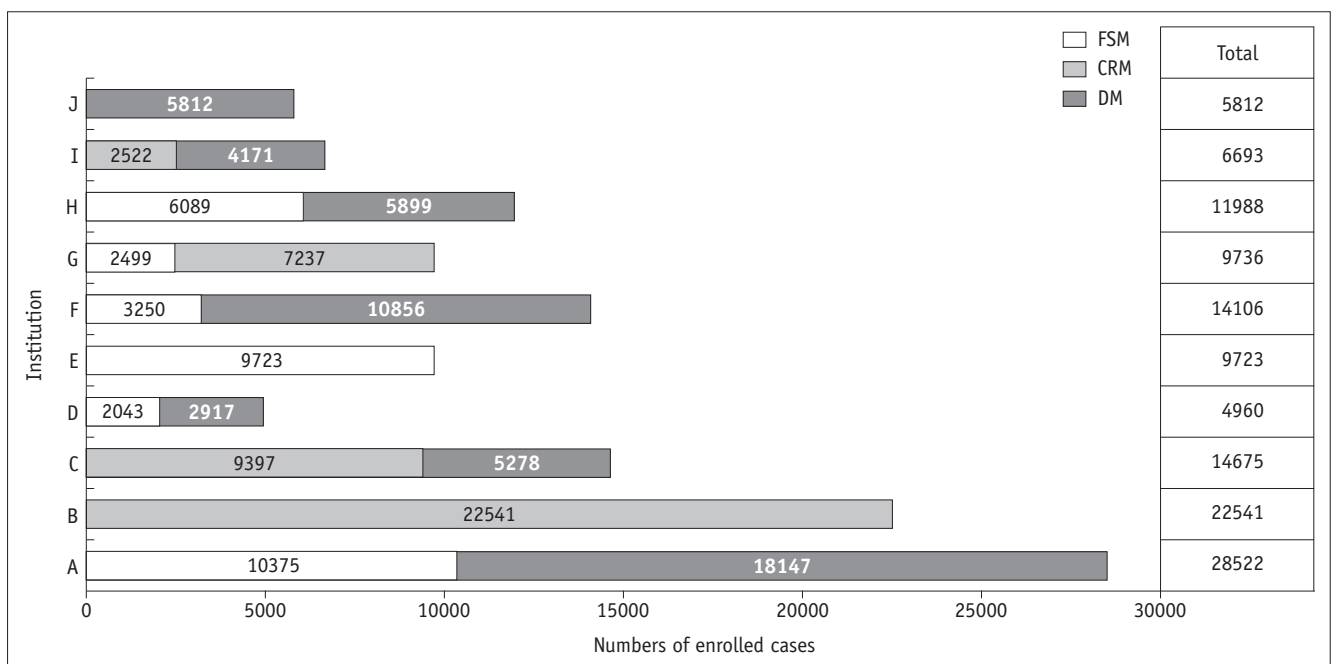


Fig. 2. Distribution of enrolled cases by institution and type of mammography equipment used in ABCS-K. Total = total number of mammography cases enrolled according to institution

FSM and CRM.

The Digital Mammography Imaging Screening Trial (DMIST) was the first study to compare the diagnostic accuracy of DM and FSM for breast cancer screening (4). In the trial, the diagnostic accuracy was similar in the entire population, which is consistent with our results. Although the AUC for DM in our study was higher than that in the DMIST (0.90 vs. 0.78), the high AUC in our study was at the cost of a higher recall rate (14.8% vs. 8.4%), and subsequently lower specificity (85.5% vs. 92.0%) and PPV₁ (2.3% vs. 5.0%).

The recall rate in our study was also higher than those

in previous studies which compared DM to FSM in Europe and the United States (6, 10, 16). The first hypothesis for the high recall rate in our study is that a high recall rate is associated with prevalence screening (17), and the high proportion (62.9%) of prevalence screening in our study could increase recall rates. Whereas, the proportions of prevalence screening in previous large-scale studies were less than 10% (6, 10). Another hypothesis for the high recall rate is that Korean women are likely to have dense breasts, and high breast density could increase recall rates (12). However, we expect that the recall rate will be further improved in the near future because our previous report revealed a downward trend in the recall rate throughout the study period (3). In addition, a study from the Netherlands showed that the recall rate dropped rapidly in the DM screening group (10). These trends of improved recall rate might be due to a better understanding of the DM findings.

The results on the recall rate, specificity, and PPV₁ were different between DM and FSM. Previous large-scale studies reported similar or inferior diagnostic performances in terms of recall rate, specificity, and PPV₁ when DM was compared with FSM (4, 6, 10). In contrast, our study showed that DM improved the recall rate, PPV₁, specificity, and FPR compared to FSM and CRM. Our results are consistent with those of a recent study, which compared the performance of DM and FSM in community practice. The aforementioned study showed improved recall rates and specificities with DM (18). These inconsistencies on the recall rate, PPV₁, and

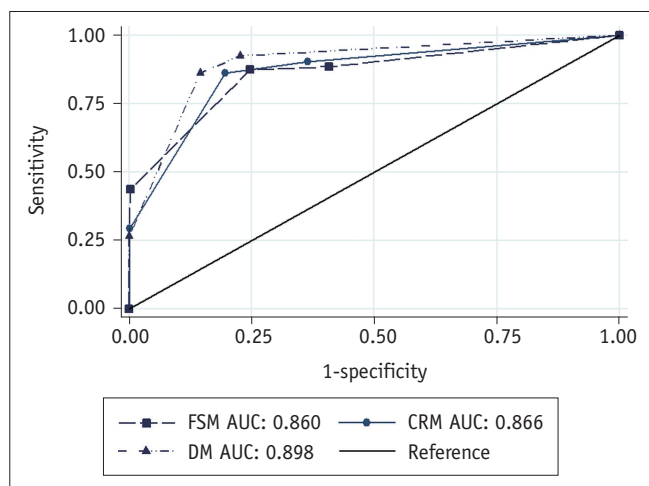


Fig. 3. Comparison of AUC for diagnostic accuracy according to type of mammography equipment used in ABCS-K. AUC = area under curve

Table 2. Diagnostic Performance of Mammography Screenings according to Types of Mammography Equipment Used

	FSM	CRM	DM	P*
Examination no. (%)	33979 (26.4)	41697 (32.4)	53080 (41.2)	
Recall rate (%)	8440/33979 (24.8)	8262/41697 (19.8)	7874/53080 (14.8)	< 0.001
CDR	2.2 (76/33979)	2.1 (88/41697)	3.4 (182/53080)	< 0.001
PPV ₁ (%)	76/8440 (0.9)	88/8262 (1.1)	182/7874 (2.3)	< 0.001
Sensitivity (%)	76/87 (87.4)	88/102 (86.3)	182/211 (86.3)	0.819
Specificity (%)	25528/33892 (75.3)	33421/41595 (80.3)	45177/52869 (85.5)	< 0.001
FPR (%)	8364/33892 (24.7)	8174/41595 (19.7)	7692/52869 (14.5)	< 0.001
ICR	0.4 (11/25539)	0.4 (14/33435)	0.6 (29/45206)	0.187

*Adjusted for readers. CDR = cancer detection rate per 1000 examinations, FPR = false positive rate, ICR = interval cancer rate per 1000 negative examinations, PPV₁ = positive predictive value₁

Table 3. Types of Breast Cancer Pathologies according to Types of Mammography Equipment Used

	Total	FSM	CRM	DM	P*
No. of all cancer patients	400	87	102	211	
No. of patients having screening-detected cancer (%)	346 (86.5)	76 (87.4)	88 (86.3)	182 (86.3)	< 0.001
Invasive cancer (%)	275 (79.5)	67 (88.2)	76 (86.4)	132 (72.5)	0.003
Carcinoma <i>in situ</i> (%)	71 (20.5)	9 (11.8)	12 (13.6)	50 (27.5)	0.003

*Adjusted for readers. Total = total numbers of patients diagnosed as having breast cancer within one year of screening

specificity of DM across studies may be due to the level of experience of radiologists, geographic regions, incidence, and study designs.

The previous results of the ABCS-K showed that the recall rate, PPV₁, specificity, and FPR of mammography screenings improved significantly from 2005 to 2010 (3). These trends in performance improvement could be partly due to the transition from the use of FSM and CRM to DM during the study period. However, in our previous report, performance significantly improved even in institutions that used only or mostly one type of mammography equipment. Therefore, various factors other than mammography equipment also likely affected performance improvement.

In our study, significantly more CISs were found when DM was used than when other equipments were used. The proportion of CIS among cancers detected with DM was almost double that detected with FSM and CRM. These findings are consistent with the results of other studies (9, 16, 19-21). Because approximately 90% of CISs are accompanied by microcalcifications (22), the high detection of CIS with DM may be related to the better visualization of microcalcifications with DM than with FSM. A recent study, which analyzed the association between detection of CIS at screening and invasive interval cancers subsequent to the screening, reported that detection and treatment of CIS is important for the prevention of future invasive cancers (23). However, a high detection of CIS with DM may not result in the anticipated reduction in cancer mortality (24-26). Therefore, long-term follow-up of the cohort in our study is required to evaluate the overdiagnosis issue of CISs. Due to recent breakthroughs in artificial intelligence, we believe that combining artificial intelligence and DM can help regulate the magnitude of overdiagnosis of CIS through helpful decision-making systems. Perhaps these issues will be the topics of future studies.

Our study has several limitations. First, we did not evaluate the association between the types of mammography equipment used and the participant factors, such as age or breast density, due to limited resources. Although the reports from the DMIST and the Breast Cancer Surveillance Consortium noted that DM was more accurate in pre- or perimenopausal women younger than 50 years with mammographically dense breasts, no such studies from Asia have been published. Therefore, further study among Asian women is needed to determine whether there are differences in the diagnostic performance of DM according to factors of the participants. Second, the

mammography equipment used in our study included two different manufacturers of DM and only one manufacturer of CRM. However, the market share of these manufacturers is dominant in Korea, and therefore, our study reflects the real world. Finally, all the participating hospitals were university-affiliated and may not be representative of the NCSP. However, the participating hospitals were distributed evenly throughout South Korea, and the proportion of DM equipment by province did not vary greatly. Therefore, these potential selection biases might have little influence on our results.

Even though our study is retrospective, it is the first multicenter study to compare the diagnostic accuracy between the types of mammography equipment used in the nationwide breast cancer screening program during the transition from the use of FSM and CRM to DM in Asia. Furthermore, unlike the previous studies that evaluated the performance of DM in which both CRM and DM were considered as DM (4), we evaluated the performance of DM and CRM separately and obtained different results.

We conclude that DM showed better performance in terms of recall rate, CDR, PPV₁, and specificity and detected more CIS than FSM and CRM. The application of DM appears to be helpful in improving the quality of mammography screenings. However, a long-term follow-up is needed to evaluate the overdiagnosis issue of CIS using DM and to determine whether breast cancer screening with DM is actually effective in reducing mortality.

Conflicts of Interest

The authors have no potential conflicts of interest to disclose.

ORCID iDs

Eun Hye Lee

<https://orcid.org/0000-0002-8773-700X>

Bo Hwa Choi

<https://orcid.org/0000-0001-7276-1579>

Jae Kwan Jun

<https://orcid.org/0000-0003-1647-0675>

Keum Won Kim

<https://orcid.org/0000-0002-7312-5483>

Young Mi Park

<https://orcid.org/0000-0001-7332-3853>

Hye-Won Kim

<https://orcid.org/0000-0002-8969-3047>

You Me Kim

<https://orcid.org/0000-0001-5807-9012>

Dong Rock Shin

<https://orcid.org/0000-0003-0722-1327>

Hyo Soon Lim

<https://orcid.org/0000-0001-6742-499X>

Jeong Seon Park

<https://orcid.org/0000-0002-5550-6645>

Hye Jung Kim

<https://orcid.org/0000-0002-0263-0941>

REFERENCES

- Jung KW, Won YJ, Oh CM, Kong HJ, Lee DH, Lee KH; Community of Population-Based Regional Cancer Registries. Cancer statistics in Korea: incidence, mortality, survival, and prevalence in 2014. *Cancer Res Treat* 2017;49:292-305
- Suh M, Choi KS, Park B, Lee YY, Jun JK, Lee DH, et al. Trends in cancer screening rates among Korean men and women: results of the Korean national cancer screening survey, 2004-2013. *Cancer Res Treat* 2016;48:1-10
- Lee EH, Kim KW, Kim YJ, Shin DR, Park YM, Lim HS, et al. Performance of screening mammography: a report of the Alliance for Breast Cancer Screening in Korea. *Korean J Radiol* 2016;17:489-496
- Pisano ED, Gatsonis C, Hendrick E, Yaffe M, Baum JK, Acharyya S, et al.; Digital Mammographic Imaging Screening Trial (DMIST) Investigators Group. Diagnostic performance of digital versus film mammography for breast-cancer screening. *N Engl J Med* 2005;353:1773-1783
- Pisano ED, Hendrick RE, Yaffe MJ, Baum JK, Acharyya S, Cormack JB, et al.; DMIST Investigators Group. Diagnostic accuracy of digital versus film mammography: exploratory analysis of selected population subgroups in DMIST. *Radiology* 2008;246:376-383
- Kerlikowske K, Hubbard RA, Miglioretti DL, Geller BM, Yankaskas BC, Lehman CD, et al.; Breast Cancer Surveillance Consortium. Comparative effectiveness of digital versus film-screen mammography in community practice in the United States: a cohort study. *Ann Intern Med* 2011;155:493-502
- Timmers JM, den Heeten GJ, Adang EM, Otten JD, Verbeek AL, Broeders MJ. Dutch digital breast cancer screening: implications for breast cancer care. *Eur J Public Health* 2012;22:925-929
- Nederend J, Duijm LE, Louwman MW, Groenewoud JH, Donkers-van Rossum AB, Voogd AC. Impact of transition from analog screening mammography to digital screening mammography on screening outcome in the Netherlands: a population-based study. *Ann Oncol* 2012;23:3098-3103
- Vigeland E, Klaasen H, Klingens TA, Hofvind S, Skaane P. Full-field digital mammography compared to screen film mammography in the prevalent round of a population-based screening programme: the Vestfold county study. *Eur Radiol* 2008;18:183-191
- van Luijt PA, Fracheboud J, Heijnsdijk EA, den Heeten GJ, de Koning HJ; National Evaluation Team for Breast Cancer Screening in Netherlands Study Group (NETB). Nation-wide data on screening performance during the transition to digital mammography: observations in 6 million screens. *Eur J Cancer* 2013;49:3517-3525
- Bae JM, Shin SY, Kim EH, Kim YN, Nam CM. Distribution of dense breasts using screening mammography in Korean women: a retrospective observational study. *Epidemiol Health* 2014;36:e2014027
- Stomper PC, D'Souza DJ, DiNitto PA, Arredondo MA. Analysis of parenchymal density on mammograms in 1353 women 25-79 years old. *AJR Am J Roentgenol* 1996;167:1261-1265
- Mandelson MT, Oestreicher N, Porter PL, White D, Finder CA, Taplin SH, et al. Breast density as a predictor of mammographic detection: comparison of interval- and screen-detected cancers. *J Natl Cancer Inst* 2000;92:1081-1087
- Sickles EA, D'Orsi CJ. *ACR BI-RADS follow-up and outcome monitoring*. In: D'Orsi CJ, Sickles EA, Mendelson EB, Morris EA, eds. *ACR BI-RADS® atlas*, 5th ed. Reston, VA: American College of Radiology, 2013:15-20
- DeLong ER, DeLong DM, Clarke-Pearson DL. Comparing the areas under two or more correlated receiver operating characteristic curves: a nonparametric approach. *Biometrics* 1988;44:837-845
- Skaane P, Hofvind S, Skjennald A. Randomized trial of screen-film versus full-field digital mammography with soft-copy reading in population-based screening program: follow-up and final results of Oslo II study. *Radiology* 2007;244:708-717
- Kim YJ, Lee EH, Jun JK, Shin DR, Park YM, Kim HW, et al.; Alliance for Breast Cancer Screening in Korea (ABCS-K). Analysis of participant factors that affect the diagnostic performance of screening mammography: a report of the Alliance for Breast Cancer Screening in Korea. *Korean J Radiol* 2017;18:624-631
- Dabbous F, Dolecek TA, Friedewald SM, Tossas-Milligan KY, Macarol T, Summerfelt WT, et al. Performance characteristics of digital vs film screen mammography in community practice. *Breast J* 2018;24:369-372
- Karssemeijer N, Bluekens AM, Beijerinck D, Deurenberg JJ, Beekman M, Visser R, et al. Breast cancer screening results 5 years after introduction of digital mammography in a population-based screening program. *Radiology* 2009;253:353-358
- Del Turco MR, Mantellini P, Ciatto S, Bonardi R, Martinelli F, Lazzari B, et al. Full-field digital versus screen-film mammography: comparative accuracy in concurrent screening cohorts. *AJR Am J Roentgenol* 2007;189:860-866
- Hedden B, Rönnow K, Olsson M, Miller D. Digital versus screen-film mammography: a retrospective comparison in a population-based screening program. *Eur J Radiol* 2007;64:419-425

22. Dershaw DD, Abramson A, Kinne DW. Ductal carcinoma in situ: mammographic findings and clinical implications. *Radiology* 1989;170:411-415
23. Duffy SW, Dibden A, Michalopoulos D, Offman J, Parmar D, Jenkins J, et al. Screen detection of ductal carcinoma in situ and subsequent incidence of invasive interval breast cancers: a retrospective population-based study. *Lancet Oncol* 2016;17:109-114
24. Duffy SW, Tabar L, Vitak B, Day NE, Smith RA, Chen HH, et al. The relative contributions of screen-detected *in situ* and invasive breast carcinomas in reducing mortality from the disease. *Eur J Cancer* 2003;39:1755-1760
25. Kalager M, Zelen M, Langmark F, Adami HO. Effect of screening mammography on breast-cancer mortality in Norway. *N Engl J Med* 2010;363:1203-1210
26. Welch HG, Prorok PC, O'Malley AJ, Kramer BS. Breast-cancer tumor size, overdiagnosis, and mammography screening effectiveness. *N Engl J Med* 2016;375:1438-1447