



Analysis of Participant Factors That Affect the Diagnostic Performance of Screening Mammography: A Report of the Alliance for Breast Cancer Screening in Korea

Young Joong Kim, MD¹, Eun Hye Lee, MD, PhD², Jae Kwan Jun, MD, PhD³, Dong-Rock Shin, MD⁴, Young Mi Park, MD, PhD⁵, Hye-Won Kim, MD, PhD⁶, Youme Kim, MD, PhD⁷, Keum Won Kim, MD, PhD¹, Hyo Soon Lim, MD, PhD⁸, Jeong Seon Park, MD, PhD⁹, Hye Jung Kim, MD¹⁰, Hye-Mi Jo, MPH³; for the Alliance for Breast Cancer Screening in Korea (ABCS-K)

¹Department of Radiology, Konyang University Hospital, Konyang University College of Medicine, Daejeon 35365, Korea; ²Department of Radiology, Bucheon Hospital, Soonchunhyang University College of Medicine, Bucheon 14584, Korea; ³National Cancer Control Institute, National Cancer Center, Goyang 10408, Korea; ⁴Department of Radiology, Gangneung Asan Hospital, University of Ulsan College of Medicine, Gangneung 25440, Korea; ⁵Department of Radiology, Busan Paik Hospital, Inje University College of Medicine, Busan 47392, Korea; ⁶Department of Radiology, Wonkwang University Hospital, Wonkwang University School of Medicine, Iksan 54538, Korea; ⁷Department of Radiology, Dankook University Hospital, Dankook University College of Medicine, Cheonan 31116, Korea; ⁸Department of Radiology, Chonnam National University Hwasun Hospital, Chonnam National University College of Medicine, Hwasun 58128, Korea; ⁹Department of Radiology, Hanyang University Hospital, Hanyang University College of Medicine, Seoul 04763, Korea; ¹⁰Department of Radiology, Kyungpook National University Medical Center, Kyungpook National University College of Medicine, Daegu 41404, Korea

Objective: To analyze participant factors that affect the diagnostic performance of screening mammography.

Materials and Methods: We enrolled 128756 cases from 10 hospitals between 2005 and 2010. We analyzed recall rate, cancer detection rate (CDR) per 1000 examinations, positive predictive value (PPV), sensitivity, specificity, false positive rate (FPR), and interval cancer rate (ICR) per 1000 negative examinations according to participant factors including age, breast density, and number of visit to the same institution, and adjusted for confounding variables.

Results: Increasing age improved recall rates (27.4% in 40's, 17.5% in 50's, 11.1% in 60's, and 8.6% in 70's), CDR (2.7, 3.2, 2.0, and 2.4), PPV (1.0, 1.8, 1.8, and 2.8%), sensitivity (81.3, 88.8, 90.3, and 94.7%), specificity (72.7, 82.7, 89.0, and 91.7%), and FPR (27.3, 17.3, 11.0, and 8.4%) ($p < 0.05$). Higher breast density impaired recall rates (4.0% in P1, 9.0% in P2, 28.9% in P3, and 27.8% in P4), PPV (3.3, 2.3, 1.2, and 1.3%), specificity (96.1, 91.2, 71.4, and 72.5%), and FPR (3.9, 8.9, 28.6, and 27.6%) ($p < 0.001$). It also increased CDR (1.3, 2.1, 3.3, and 3.6) and ICR (0.2, 0.3, 0.6, and 1.6) ($p < 0.05$). Successive visits to the same institution improved recall rates (20.9% for one visit, 10.7% for two visits, 7.7% for more than three visits), PPV (1.6, 2.8, and 2.7%), specificity (79.4, 89.6, and 92.5%), and FPR (20.6, 10.4, and 7.5%) ($p < 0.001$).

Conclusion: Young age and dense breasts negatively affected diagnostic performance in mammography screening, whereas successive visits to the same institution had a positive effect. Examinee education for successive visits to the same institution would improve the diagnostic performance.

Keywords: Screening mammography; Cancer detection rates; Audit; National mammography database

Received October 19, 2016; accepted after revision December 27, 2016.

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, 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 (<http://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

Screening mammography is useful for detecting early breast cancer and reducing breast cancer mortality (1). Assessment of the diagnostic performance of screening mammography has become increasingly important to promote cancer prevention and improve the quality of screening programs. The Breast Cancer Surveillance Consortium (BCSC) is a research initiative funded by the National Cancer Institute that incorporates seven mammography registries, and features a Statistical Coordinating Center that collects and analyzes mammography and pathology data in the United States. The BCSC is a valuable resource when it is sought to assess the delivery and quality of breast cancer screening and patient-related outcomes (2). Several studies have reported the diagnostic performance for screening mammography from BCSC data, including recall rates of 4.0–19.6%, sensitivity 73.0–95.0%, specificity 83.0–98.5%, positive predictive value (PPV) 2.6–8.0%, and cancer detection rate (CDR) 2.0–4.7 per 1000 examinations (3–8). They also provided desirable goals and the threshold of minimally acceptable interpretive performance for screening mammography (4, 6, 9).

In Korea, women age 40 years or older have undergone biannual mammography screening under the auspices of the National Cancer Screening Program (NCSP) since 1999 (10). However, no database allowing assessment of the quality of such screening has yet been established. Therefore, we organized the Alliance for Breast Cancer Screening in Korea (ABCS-K); 11 institutions agreed to analyze and report on the diagnostic performance of screening mammography (11). It was recently shown that, between 2005 and 2010, the sensitivity and CDR of screening mammography among ABCS-K institutions were comparable to those of Western programs. However, the recall rate, PPV, and specificity were suboptimal despite significant improvements during the study period.

Some authors have identified participant factors affecting the diagnostic performance of screening mammography (5, 7, 8, 12, 13), but no such report has emanated from Korea. In the present study, we explored participant factors affecting the diagnostic performance of screening mammography in Korea; we used the ABCS-K data to this end.

MATERIALS AND METHODS

Study Population

Our study was approved by the Institutional Review Board

of all the participating hospitals; the need for informed participant consent was waived.

We collected data from participating hospitals between January 2005 and December 2010; the data were linked with the NCSP database. We excluded mammographic data on women with histories of breast cancer, those who received interstitial injections, and those who could not be reliably identified. We enrolled 128756 women examined in 10 hospitals over 6 years. The study population and data collection methods were described in our previous report (11).

Screening Mammography

We obtained two mammographic images, in the mediolateral oblique and craniocaudal planes. Each mammographic report contained a visual assessment of breast density as defined by the American College of Radiology (ACR) Breast Imaging-Reporting and Data System, on a scale of 1 to 4. Pattern 4 (P4) is “extremely dense” breast tissue that may lower the sensitivity of mammography; pattern 3 (P3) is “heterogeneously dense” tissue that may prevent detection of small masses; pattern 2 (P2) features “scattered fibroglandular densities”; and pattern 1 (P1) describes breasts that are “almost entirely fat” (14). Breasts of patterns 1 and 2 were considered to be fatty breasts, and those of patterns 3 and 4 were considered to be dense breasts.

Outcome Measurements and Statistical Analysis

We used the National Health Insurance Service (NHIS) database. The data were linked with the NCSP database and National Cancer Center registry for breast cancer. We collected information on participants, radiologists, and facilities with the approval of the Institutional Review Boards of all participating hospitals. These were merged into the NHIS database.

We calculated performance indicators including the recall rate, CDR per 1000 examinations, PPV, sensitivity, specificity, false-positive rate (FPR), and interval cancer rate (ICR) per 1000 negative examinations. The performance indicators were defined in our previous report (11). We analyzed all data by age, breast density, and the number of visits to the same institution; we calculated odds ratios (ORs) (both crude [cORs] and adjusted [aORs], with 95% confidence intervals). The OR reference values were those for women aged 40–49 years, with breasts of density P1, and who visited screening institutions only once. Each performance indicator was adjusted for other factors. Of the

Table 1. Distribution of Participants' Breast Density and Number of Visits to the same Institution by Age Group

	Age Group				P for Trend*
	40-49	50-59	60-69	≥ 70	
Breast density					
Total, n [†]	45449	47356	28331	7620	< 0.001
P1 (%)	1.7	10.3	30.2	52.1	
P2 (%)	15.2	36.9	45.3	37.1	
P3 (%)	64.4	47.5	23.7	10.6	
P4 (%)	18.8	5.3	0.9	0.2	
Number of visits					
Total, n [‡]	36728	37225	22774	6684	< 0.001
Once (%)	86.1	77.0	75.2	78.1	
Twice (%)	11.6	17.9	19.2	17.0	
≥ 3 times	2.3	5.0	5.6	4.9	

*p value of chi-square test, [†]total number of examinations, [‡]total number of examinees. P1 = almost entirely fatty, P2 = scattered fibroglandular density, P3 = heterogeneously dense, P4 = extremely dense

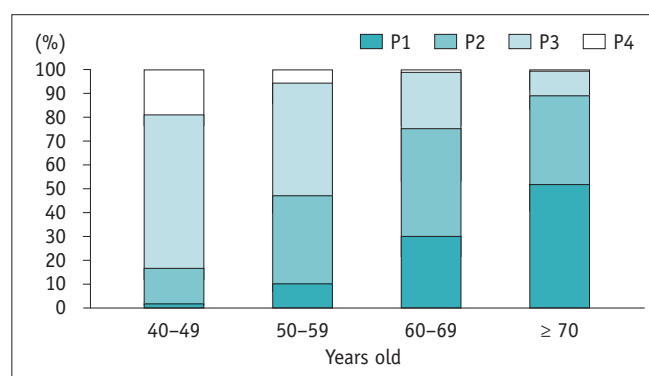


Fig. 1. Final breast density classifications by age groups.

Following BI-RADS criteria, breast density was classified as P1, almost entirely fatty; P2, scattered fibroglandular density; P3, heterogeneously dense; P4, extremely dense. BI-RADS = Breast Imaging-Reporting and Data System

cORs and aORs, the latter were more meaningful because possible confounding factors were controlled.

We also analyzed trends among indicators evident among the various grades; a p value of less than 0.05 was considered significant. We used logistic regression to compare differences among grades, and the trends afforded by the indicators. We used SAS ver. 9.2 (SAS Institute, Cary, NC, USA), to this end.

RESULTS

Demographic Factors

Table 1 shows breast densities and the numbers of visits to the same institution by age group. Supplementary Table 1 (in the online-only Data Supplement) shows the distribution of the enrolled participants by participating hospital. Breast density differed significantly by age group;

the proportion of dense breasts decreased significantly with increasing age (Table 1, Fig. 1). The number of visits to the same institution differed significantly by age group, increasing significantly with advanced age (Table 1).

Diagnostic Performance

Table 2 shows the diagnostic performance of screening mammography by age group. The recall rates of older women were significantly lower than that of women in their 40s. The recall rate decreased significantly with increasing age. The CDR of women in their 50s was highest among all age groups, but the CDR improved with increasing age after adjusting for breast density and number of visits to the same institution. The PPVs of women in their 50s and 70s were higher than those of women in their 40s, but the PPVs of women in their 60s did not differ. The trend in PPV was upward with increasing age. The sensitivities did not differ among the age groups, but trended significantly upward with increasing age. The specificities of older women were higher than that of women in their 40s and trended upward with increasing age. The FPRs of older women were lower than that of women in their 40s and trended downward with increasing age. The ICR did not differ among the age groups and did not trend up or down with increasing age after the adjustment.

Table 3 shows the diagnostic performance of screening mammography by breast density. The recall rates of higher density groups were significantly higher than that of the P1 group. The rate trended significantly upward with increasing breast density. The CDRs of higher density women were greater than that of the P1 group and trended upward with increasing breast density. The PPVs of women with P3

Table 2. Diagnostic Performance of Screening Mammography According to Age Group

	Age group				P for Trend
	40-49	50-59	60-69	≥ 70	
Recall rate (%)	27.4 (27.0-27.9)	17.5 (17.2-17.9)	11.1 (10.8-11.5)	8.6 (7.9-9.2)	
cOR	1.00 (ref.)	0.56 (0.55-0.58)	0.33 (0.32-0.35)	0.25 (0.23-0.27)	< 0.001
aOR*	1.00 (ref.)	0.80 (0.78-0.83)	0.76 (0.72-0.79)	0.83 (0.76-0.91)	< 0.001
CDR per 1000 examinations	2.7 (2.2-3.2)	3.2 (2.7-3.7)	2.0 (1.5-2.5)	2.4 (1.3-3.5)	
cOR	1.00 (ref.)	1.18 (0.93-1.50)	0.74 (0.54-1.01)	0.88 (0.54-1.44)	0.136
aOR*	1.00 (ref.)	1.52 (1.18-1.95)	1.21 (0.86-1.72)	1.68 (0.98-2.88)	0.037
PPV (%)	1.0 (0.8-1.2)	1.8 (1.5-2.1)	1.8 (1.3-2.2)	2.8 (1.5-4.0)	
cOR	1.00 (ref.)	1.86 (1.47-2.37)	1.83 (1.33-2.52)	2.87 (1.74-4.74)	< 0.001
aOR*	1.00 (ref.)	1.75 (1.36-2.25)	1.39 (0.97-1.99)	1.87 (1.08-3.24)	0.003
Sensitivity (%)	81.3 (75.1-87.6)	88.8 (84.0-93.5)	90.3 (83.0-97.7)	94.7 (84.7-100.0)	
cOR	1.00 (ref.)	1.81 (0.97-3.40)	2.14 (0.84-5.47)	4.13 (0.53-32.21)	0.022
aOR*	1.00 (ref.)	1.73 (0.89-3.37)	2.54 (0.83-7.78)	5.08 (0.58-44.46)	0.028
Specificity (%)	72.7 (72.3-73.2)	82.7 (82.4-83.1)	89.0 (88.7-89.4)	91.7 (91.0-92.3)	
cOR	1.00 (ref.)	1.80 (1.74-1.85)	3.04 (2.92-3.18)	4.11 (3.78-4.47)	< 0.001
aOR*	1.00 (ref.)	1.26 (1.22-1.30)	1.33 (1.27-1.39)	1.22 (1.11-1.34)	< 0.001
FPR (%)	27.3 (26.9-27.7)	17.3 (16.9-17.6)	11.0 (10.6-11.3)	8.4 (7.7-9.0)	
cOR	1.00 (ref.)	0.56 (0.54-0.58)	0.33 (0.32-0.34)	0.24 (0.22-0.27)	< 0.001
aOR*	1.00 (ref.)	0.80 (0.77-0.82)	0.75 (0.72-0.79)	0.82 (0.75-0.90)	< 0.001
ICR per 1000 negative examinations	0.8 (0.5-1.2)	0.5 (0.3-0.7)	0.2 (0.0-0.4)	0.1 (0.0-0.4)	
cOR	1.00 (ref.)	0.57 (0.32-1.03)	0.28 (0.12-0.68)	0.17 (0.02-1.24)	< 0.001
aOR*	1.00 (ref.)	0.83 (0.45-1.55)	0.50 (0.19-1.34)	0.32 (0.04-2.57)	0.118

*Adjusted with breast density and number of visits to same institution, parentheses = 95% confidence interval. aOR = adjusted odds ratio, CDR = cancer detection rate, cOR = crude odds ratio, FPR = false positive rate, ICR = interval cancer rate, PPV = positive predictive value

Table 3. Diagnostic Performance of Screening Mammography According to Breast Density

	P1	P2	P3	P4	P for Trend
	Recall rate (%)	4.0 (3.7-4.3)	9.0 (8.8-9.3)	28.9 (28.5-29.2)	
cOR	1.00 (ref.)	2.39 (2.21-2.60)	9.77 (9.05-10.55)	9.27 (8.51-10.09)	< 0.001
aOR*	1.00 (ref.)	2.32 (2.13-2.52)	8.70 (8.02-9.43)	7.69 (7.02-8.43)	< 0.001
CDR per 1000 examinations	1.3 (0.8-1.9)	2.1 (1.6-2.5)	3.3 (2.9-3.8)	3.6 (2.5-4.7)	
cOR	1.00 (ref.)	1.57 (1.00-2.47)	2.53 (1.66-3.87)	2.75 (1.66-4.56)	< 0.001
aOR*	1.00 (ref.)	1.64 (1.03-2.61)	2.83 (1.79-4.49)	3.33 (1.91-5.79)	< 0.001
PPV (%)	3.3 (2.0-4.6)	2.3 (1.8-2.8)	1.2 (1.0-1.3)	1.3 (0.9-1.7)	
cOR	1.00 (ref.)	0.68 (0.43-1.08)	0.34 (0.22-0.53)	0.39 (0.23-0.64)	< 0.001
aOR*	1.00 (ref.)	0.70 (0.44-1.13)	0.41 (0.26-0.66)	0.51 (0.29-0.90)	0.001
Sensitivity (%)	85.7 (72.8-98.7)	88.3 (81.8-94.8)	88.4 (84.2-92.6)	75.9 (64.5-87.3)	
cOR	1.00 (ref.)	1.26 (0.37-4.31)	1.27 (0.41-3.95)	0.53 (0.15-1.80)	0.187
aOR*	1.00 (ref.)	1.58 (0.44-5.75)	2.26 (0.61-8.33)	1.08 (0.26-4.49)	0.821
Specificity (%)	96.1 (95.9-96.4)	91.2 (90.9-91.4)	71.4 (71.0-71.7)	72.5 (71.6-73.3)	
cOR	1.00 (ref.)	0.41 (0.38-0.45)	0.10 (0.09-0.11)	0.11 (0.10-0.12)	< 0.001
aOR*	1.00 (ref.)	0.43 (0.39-0.47)	0.11 (0.10-0.12)	0.13 (0.12-0.14)	< 0.001
FPR (%)	3.9 (3.6-4.1)	8.9 (8.6-9.1)	28.6 (28.3-29.0)	27.6 (26.7-28.4)	
cOR	1.00 (ref.)	2.42 (2.23-2.63)	9.99 (9.24-10.8)	9.47 (8.69-10.33)	< 0.001
aOR*	1.00 (ref.)	2.34 (2.15-2.55)	8.86 (8.16-9.62)	7.82 (7.13-8.58)	< 0.001
ICR per 1000 negative examinations	0.2 (0.0-0.5)	0.3 (0.1-0.5)	0.6 (0.4-0.9)	1.6 (0.7-2.5)	
cOR	1.00 (ref.)	1.32 (0.42-4.13)	2.69 (0.94-7.70)	6.96 (2.27-21.34)	< 0.001
aOR*	1.00 (ref.)	1.05 (0.32-3.40)	1.75 (0.56-5.54)	4.04 (1.15-14.27)	0.005

*Adjusted with age and number of visits to same institution, parentheses = 95% confidence interval. aOR = adjusted odds ratio, CDR = cancer detection rate, cOR = crude odds ratio, FPR = false positive rate, ICR = interval cancer rate, PPV = positive predictive value, P1 = almost entirely fatty, P2 = scattered fibroglandular density, P3 = heterogeneously dense, P4 = extremely dense

and P4 breasts were lower than those of P1 women, but the PPVs of those with P2 did not differ from those of P1 women. The trend was downward with increasing density. The sensitivities of the density groups did not significantly differ and no trend up or down was evident as breast density increased. The specificities of the higher density groups were lower than that of the P1 group and trended downward with increasing density. The FPRs of the higher density groups were greater than that of the P1 group and trended upward with increasing density. The ICR of only the P4 group was higher than that of the P1 group, but values generally trended upward with increasing breast density.

The diagnostic performances upon division of women into two groups by breast density (fatty and dense) are shown in Supplementary Table 2 (in the online-only Data Supplement). The OR reference standard was based on data from women with fatty breasts (grades P1 and P2). The recall rates, CDRs, FPR, and ICRs were significantly higher in women with dense breasts after adjusting for age and the number of visits. The PPVs and specificities were lower among women with dense breasts. The sensitivities did not differ significantly between the groups, even after the adjustment.

Table 4 shows the diagnostic performance of screening mammography according to the number of visits to the same institution. The recall rates were significantly lower after successive visits than after one visit. The recall rate trended significantly downward with increasing numbers of visits. The PPVs associated with two visits were higher than those associated with one visit. The PPV generally trended upward as the number of visits increased. The specificities associated with successive visits were higher than that associated with one visit, and trended upward with increasing numbers of visits. The FPR of women who visited repeatedly was lower than that of one-visit women. None of CDR, sensitivity, or ICR differed by numbers of visits; no trend was evident.

DISCUSSION

We recently reported on the diagnostic performance of, and annual trends in, screening mammography conducted from 2005 to 2010 (11). Compared with the desirable goals of the ACR and benchmarks of the BCSC, the sensitivity (86.5%) and CDR (2.7) were comparable, while the

Table 4. Diagnostic Performance of Screening Mammography According to Number of Visits to the same Institution

	Number of Visits			P for Trend
	Once	Twice	≥ 3 Times	
Recall rate (%)	20.9 (20.6–21.2)	10.7 (10.2–11.2)	7.7 (6.9–8.5)	
cOR	1.00 (ref.)	0.45 (0.43–0.48)	0.32 (0.28–0.35)	< 0.001
aOR*	1.00 (ref.)	0.48 (0.46–0.51)	0.35 (0.31–0.39)	< 0.001
CDR per 1000 examinations	3.3 (2.9–3.6)	3.0 (2.1–3.8)	2.1 (0.7–3.4)	
cOR	1.00 (ref.)	0.92 (0.68–1.24)	0.64 (0.33–1.24)	0.187
aOR*	1.00 (ref.)	0.92 (0.68–1.25)	0.64 (0.33–1.24)	0.193
PPV (%)	1.6 (1.4–1.7)	2.8 (2.0–3.6)	2.7 (1.0–4.5)	
cOR	1.00 (ref.)	1.81 (1.33–2.47)	1.76 (0.90–3.45)	< 0.001
aOR*	1.00 (ref.)	1.65 (1.21–2.26)	1.55 (0.79–3.06)	< 0.001
Sensitivity (%)	86.8 (83.0–90.6)	83.1 (73.5–92.6)	90.0 (71.4–100.0)	
cOR	1.00 (ref.)	0.75 (0.35–1.59)	1.37 (0.17–11.10)	0.711
aOR*	1.00 (ref.)	0.79 (0.36–1.72)	1.24 (0.15–10.61)	0.741
Specificity (%)	79.4 (79.1–79.7)	89.6 (89.1–90.1)	92.5 (91.7–93.3)	
cOR	1.00 (ref.)	2.23 (2.12–2.36)	3.21 (2.86–3.60)	< 0.001
aOR*	1.00 (ref.)	2.09 (1.98–2.21)	2.89 (2.57–3.25)	< 0.001
FPR (%)	20.6 (20.4–20.9)	10.4 (10.0–10.9)	7.5 (6.7–8.3)	
cOR	1.00 (ref.)	0.45 (0.42–0.47)	0.31 (0.28–0.35)	< 0.001
aOR*	1.00 (ref.)	0.48 (0.45–0.51)	0.35 (0.31–0.39)	< 0.001
ICR per 1000 negative examinations	0.6 (0.4–0.8)	0.7 (0.3–1.1)	0.3 (0.0–0.7)	
cOR	1.00 (ref.)	1.09 (0.54–2.17)	0.40 (0.06–2.90)	0.604
aOR*	1.00 (ref.)	1.17 (0.58–2.34)	0.45 (0.06–3.29)	0.780

*Adjusted with age and breast density, parentheses = 95% confidence interval. aOR = adjusted odds ratio, CDR = cancer detection rate, cOR = crude odds ratio, FPR = false positive rate, ICR = interval cancer rate, PPV = positive predictive value

specificity (81.1%) and PPV (1.4%) were low and the recall rate (19.1%) was high (6, 14).

We suggested that many factors affect mammographic performance. In the present study, we show that participant factors including age, breast density, and the number of visits to the same institution constantly affected the recall rate, PPV, specificity, and FPR. Younger age and higher breast density were associated with a higher recall rate and FPR, and lower PPV and specificity impairing the efficacy of mammography screening. By contrast, successive visits to the same institution were associated with lower recall rate and FPR, and higher PPV and specificity, thus improving quality of mammography screening.

In our study, the proportion of dense breasts was slightly higher than that of fatty breasts. However, it was similar to other studies (5, 8). The distribution of age was also similar to that of other studies (5, 6, 8). Women in their 40s and 50s accounted for more than half. Among them, women in their 50s constituted the greatest proportion of participants in this study. Although the incidence of breast cancer peaks in women aged in their late 40s in Korea (15), the CDR was highest among women in their 50s in the present study, probably attributable to the fact that most women studied were in their 50s.

We found that younger women had high proportions of dense breasts (Table 1, Fig. 1). This may affect mammographic performance because high breast density is a confounding factor. Therefore, to evaluate each diagnostic parameter, we adjusted for the other factors. We adjusted for breast density and the number of visits to the same institution when evaluating diagnostic performance by age; we adjusted for age and numbers of visits when evaluating performance by breast density; and we adjusted for age and breast density when evaluating performance by number of visits. This rendered the aORs more meaningful. We offer an example of the difference between cOR and aOR: the CDR trended downward with increasing age after adjustment. However, the cOR was not apparent (Table 2), whereas, the ICR trended downward with increasing age when the cOR was evaluated. Notably, after the adjustment, no significant trend was apparent (Table 2).

It has long been known that dense breast tissue may obscure breast cancers, decreasing the sensitivity of screening mammography (16-18). The lower sensitivity of mammography in younger women, especially those aged 40-49 years, may be attributable to the fact that their breasts are radiographically dense. Rapid tumor growth in younger

women may also cause more interval cancers and may be the prime explanation as to why sensitivity of screening mammography is lower in younger women (7, 19-21).

However, we found that sensitivity did not differ by age and breast density. Sensitivity generally trended upward with increasing age, but the significance of the trend was lower than that of other significant parameters. This may be explained by the very high recall rates; these were optimal only in women aged more than 70 years (8.6% vs. 11.1-27.4%) with P1 breast densities (4.0% vs. 27.8-28.9%); the proportions of those aged over 70 years and those with breasts of P1 density were minor. Thus, unnecessarily high recall rates may conceal limitations imparted by younger age and dense breasts on the effectiveness of mammography screening. Also, ICR was not affected by younger age.

It has been reported that dense breast tissue *per se* is an independent risk factor for breast cancer (22, 23). We found that the CDR and ICR increased as breast density rose, even after the adjustment (Table 3); this is in line with those of other studies (17, 22).

Successive visits to the same institution improved the quality of screening mammography. Comparison of current and prior mammograms significantly improves overall performance and reduces the recall rate (24). By contrast, the CDR trended downward as the number of visits increased, but this did not attain statistical significance. This may be associated with a lower rate of incident cancers compared with prevalent cancers.

We found that participant age, breast density, and the number of visits to the same institution affected the overall diagnostic performance, but not the sensitivity, of screening mammography. Neither age nor breast density is modifiable, but participant behavior may be changed via education, which should be imparted to women undergoing screening mammography. Several studies have suggested that relevant education programs about breast cancer screening could improve participant behavior (25-27). It is important to decrease the unnecessarily high recall rate, especially among younger women with dense breasts; this would improve the quality of screening mammography. In addition, as screening mammography has certain limitations in young women with dense breasts, supplementary ultrasound or digital tomosynthesis (28) evaluation should be explored.

Our present study had several limitations. First, we did not evaluate certain risk factors, including family history, because the study was retrospective in nature and such information was not available. Second, we encountered

some “interval” cancers, but could not differentiate false-negative from true interval cancers because the Personal Information Protection Law dictated that we were to be given only the numbers of screening-detected and interval cancers; we were denied detailed information on the cancer patients. Interval cancer does not always indicate false-negativity, although it could be regarded as a false negative in population-based screening programs (21). Third, the participant factors recorded differed somewhat by screening institution (Supplementary Table 1 in the online-only Data Supplement) (11). However, our study is meaningful because this is the first multicenter work exploring the effects of participant factors on the efficacy of the screening mammography run by the NCSP in Korea.

In conclusion, younger age and higher breast density negatively affected the diagnostic performance of screening mammography in terms of the recall rate, PPV, specificity, and FPR; successive visits to the same institution had a positive effect. If examinees were encouraged to successive visits to the same institution, this would improve the diagnostic performance of breast-cancer screening.

Supplementary Materials

The online-only Data Supplement is available with this article at <https://doi.org/10.3348/kjr.2017.18.4.624>.

REFERENCES

- Humphrey LL, Helfand M, Chan BK, Woolf SH. Breast cancer screening: a summary of the evidence for the U.S. Preventive Services Task Force. *Ann Intern Med* 2002;137(5 Part 1):347-360
- Ballard-Barbash R, Taplin SH, Yankaskas BC, Ernster VL, Rosenberg RD, Carney PA, et al. Breast Cancer Surveillance Consortium: a national mammography screening and outcomes database. *AJR Am J Roentgenol* 1997;169:1001-1008
- Elmore JG, Jackson SL, Abraham L, Miglioretti DL, Carney PA, Geller BM, et al. Variability in interpretive performance at screening mammography and radiologists' characteristics associated with accuracy. *Radiology* 2009;253:641-651
- Carney PA, Sickles EA, Monsees BS, Bassett LW, Brenner RJ, Feig SA, et al. Identifying minimally acceptable interpretive performance criteria for screening mammography. *Radiology* 2010;255:354-361
- Carney PA, Miglioretti DL, Yankaskas BC, Kerlikowske K, Rosenberg R, Rutter CM, et al. Individual and combined effects of age, breast density, and hormone replacement therapy use on the accuracy of screening mammography. *Ann Intern Med* 2003;138:168-175
- Rosenberg RD, Yankaskas BC, Abraham LA, Sickles EA, Lehman CD, Geller BM, et al. Performance benchmarks for screening mammography. *Radiology* 2006;241:55-66
- Kerlikowske K, Grady D, Barclay J, Sickles EA, Ernster V. Effect of age, breast density, and family history on the sensitivity of first screening mammography. *JAMA* 1996;276:33-38
- Kerlikowske K, Zhu W, Hubbard RA, Geller B, Dittus K, Braithwaite D, et al. Outcomes of screening mammography by frequency, breast density, and postmenopausal hormone therapy. *JAMA Intern Med* 2013;173:807-816
- Sickles EA, Wolverson DE, Dee KE. Performance parameters for screening and diagnostic mammography: specialist and general radiologists. *Radiology* 2002;224:861-869
- 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
- Kolb TM, Lichy J, Newhouse JH. Comparison of the performance of screening mammography, physical examination, and breast US and evaluation of factors that influence them: an analysis of 27,825 patient evaluations. *Radiology* 2002;225:165-175
- Laya MB, Larson EB, Taplin SH, White E. Effect of estrogen replacement therapy on the specificity and sensitivity of screening mammography. *J Natl Cancer Inst* 1996;88:643-649
- D'Orsi CJ, Bassett LW, Berg WA, Feig SA, Jackson VP, Kopans DB, et al. *BI-RADS: mammography*. In: D'Orsi CJ, Mendelson EB, Ikeda DM, eds. *BIRADS® breast imaging reporting and data system: breast imaging atlas*, 4th ed. Reston, VA: American College of Radiology, 2003
- Park SK, Kang D, Kim Y, Yoo KY. Epidemiologic characteristics of the breast cancer in Korea. *J Korean Med Assoc* 2009;52:937-945
- Murphy IG, Dillon MF, Doherty AO, McDermott EW, Kelly G, O'Higgins N, et al. Analysis of patients with false negative mammography and symptomatic breast carcinoma. *J Surg Oncol* 2007;96:457-463
- 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
- Checka CM, Chun JE, Schnabel FR, Lee J, Toth H. The relationship of mammographic density and age: implications for breast cancer screening. *AJR Am J Roentgenol* 2012;198:W292-W295
- Tabar L, Fagerberg G, Chen HH, Duffy SW, Smart CR, Gad A, et al. Efficacy of breast cancer screening by age. New results from the Swedish Two-County Trial. *Cancer* 1995;75:2507-2517
- Brekelmans CT, Collette HJ, Collette C, Fracheboud J, de Waard F. Breast cancer after a negative screen: follow-up of

- women participating in the DOM Screening Programme. *Eur J Cancer* 1992;28A:893-895
21. Suzuki A, Kuriyama S, Kawai M, Amari M, Takeda M, Ishida T, et al. Age-specific interval breast cancers in Japan: estimation of the proper sensitivity of screening using a population-based cancer registry. *Cancer Sci* 2008;99:2264-2267
 22. Boyd NF, Guo H, Martin LJ, Sun L, Stone J, Fishell E, et al. Mammographic density and the risk and detection of breast cancer. *N Engl J Med* 2007;356:227-236
 23. Harvey JA, Bovbjerg VE. Quantitative assessment of mammographic breast density: relationship with breast cancer risk. *Radiology* 2004;230:29-41
 24. Roelofs AA, Karssemeijer N, Wedekind N, Beck C, van Woudenberg S, Snoeren PR, et al. Importance of comparison of current and prior mammograms in breast cancer screening. *Radiology* 2007;242:70-77
 25. Wang WL, Hsu SD, Wang JH, Huang LC, Hsu WL. Survey of breast cancer mammography screening behaviors in Eastern Taiwan based on a health belief model. *Kaohsiung J Med Sci* 2014;30:422-427
 26. Moodi M, Rezaeian M, Mostafavi F, Sharifirad GR. Determinants of mammography screening behavior in Iranian women: a population-based study. *J Res Med Sci* 2012;17:750-759
 27. Juon HS, Kim M, Shankar S, Han W. Predictors of adherence to screening mammography among Korean American women. *Prev Med* 2004;39:474-481
 28. Kim SA, Chang JM, Cho N, Yi A, Moon WK. Characterization of breast lesions: comparison of digital breast tomosynthesis and ultrasonography. *Korean J Radiol* 2015;16:229-238