Original Article | Breast Imaging
eISSN 2005-8330
https://doi.org/10.3348/kjr.2024.0866
Korean J Radiol 2025;26(4):313-323

Korean Journal of Radiology
KJR

Biennial Mammography Performance in the Korean National Cancer Screening Program From 2009 to 2020

Dong Jun Kim[1,2], Nan-He Yoon[3], Jae Kwan Jun[4,5], Mina Suh[4,5], Sunhwa Lee[4], Seongju Kim[1], Horim A. Hwang[1], Seung Eun Jung[6], Hooyeon Lee[1,2]

[1]Department of Preventive Medicine, College of Medicine, The Catholic University of Korea, Seoul, Republic of Korea
[2]Department of Public Health, Graduate School, The Catholic University of Korea, Seoul, Republic of Korea
[3]Division of Social Welfare and Health Administration, Wonkwang University, Iksan, Republic of Korea
[4]National Cancer Control Institute, National Cancer Center, Goyang, Republic of Korea
[5]Graduate School of Cancer Science and Policy, National Cancer Center, Goyang, Republic of Korea
[6]Department of Radiology, Eunpyeong St. Mary's Hospital, College of Medicine, The Catholic University of Korea, Seoul, Republic of Korea

Objective: Mammography is essential for reducing breast cancer mortality; however, its performance varies globally. This study aimed to evaluate mammography screening outcomes in Korea over 12 years and investigate regional variations.

Materials and Methods: We analyzed mammography data from 42 million Korean women, aged 40 years and older, who participated in the Korean National Cancer Screening Program (KNCSP) from 2009 to 2020. Performance metrics—including recall rate (RR), positive predictive value (PPV), sensitivity, specificity, false positive rate (FPR), cancer detection rate (CDR), interval cancer rate (ICR), and dense breast rate (DBR), were computed. Twelve-year trends in these metrics were analyzed using Joinpoint regression. Regional variations were also examined across Korea's 237 districts, stratified by age groups.

Results: From 2009 to 2020, 42165405 mammography screenings were conducted through the KNCSP, increasing from 2821132 screenings in 2009 to 3596204 in 2020. The RR decreased from 17.2% in 2009 to 11.2% in 2020 (average annual percent change [AAPC] = -3.7%), while the PPV increased from 0.8% to 2.8%; AAPC = 10.7%), the CDR increased from 1.5 to 3.1 per 1000; AAPC = 7.3%), and the ICR rose from 0.9 to 1.6 per 1000; (AAPC = 5.2%). Regional variations were noted; however, differences in the RR, sensitivity, specificity, and FPR decreased over time.

Conclusion: While mammography performance improved from 2009 to 2020, the PPV and sensitivity remain suboptimal, underscoring the need for continuous monitoring. Regional disparities in performance, although reduced, persist. These findings provide essential baseline data for improving mammography quality and addressing inequities in breast cancer screening.

Keywords: Breast cancer; Mammography; Performance; Outcome; Result; Trend analysis; Regional variation

INTRODUCTION

Breast cancer (BC) is the predominantly diagnosed cancer worldwide, and its burden has increased over the past decades [1]. Mammography remains the primary screening method for BC and is an effective tool for reducing BC mortality [2-4]. To ensure high quality mammography, the American College of Radiology Breast Imaging Reporting and Data System (BI-RADS) and the Mammography Quality Standards Act mandate the auditing of interpretive performance in screening mammography [5,6].

In the United States (US), the National Mammography Database (NMD), established in 2008, serves as a quality improvement tool [7]. Acceptable NMD performance metrics for screening mammography include a recall rate (RR) between 5% and 12%, a cancer detection rate (CDR) of at least 2.5 per 1000 screening examinations, and a positive predictive value (PPV) between 3% and 8% [6,7]. Another national mammography screening and outcomes database, the Breast Cancer Surveillance Consortium (BCSC), suggested

Received: September 2, 2024 **Revised:** January 5, 2025
Accepted: January 6, 2025
Corresponding author: Hooyeon Lee, MD, PhD, Department of Preventive Medicine, College of Medicine, The Catholic University of Korea, 222 Banpo-daero, Seocho-gu, Seoul 06591, Republic of Korea
• E-mail: hylee@catholic.ac.kr

that the acceptable combined sensitivity and specificity ranges for screening mammography are a sensitivity ≥80% and specificity ≥85% or sensitivity ranging from 75%–79% and specificity from 88%–97%, and the ranges for CDR, RR, and PPV are as follows: a CDR ≥6 per 1000, RR from 3%–20%, and any PPV; a CDR ranging from 4–6 per 1000, RR from 3%–15%, and PPV ≥3%; or a CDR from 2.5–4.0 per 1000, RR from 5%–12%, and PPV from 3%–8% [8,9]. However, a benchmark for screening mammography specifically for Asian women is yet to be established.

BC is traditionally more prevalent among women in the west, with an understanding of this condition and its management largely based on research and data extracted from western populations [10]. However, emerging differences in the epidemiology, tumor biology, and host biology of BC in Asian populations may have implications for its clinical management in ethnically diverse populations [10]. Further research and data from larger Asian populations are required to better understand and manage BC. Moreover, as women age, the decrease in breast density and the rise in BC incidence impact the performance of mammography [10,11]. Variations in mammography performance are observed based on the geographic location of radiology practices in the US [6,10].

Since 2002, all Korean women aged ≥40 years have received biennial invitations for mammographic screening through the Korean National Cancer Screening Program (KNCSP) [12]. In contrast to the US and Europe, there is currently no consensus in Korea regarding acceptable ranges for the performance metrics of screening mammography [13]. To the best of our knowledge, studies on mammography performance in the KNCSP have either described the early period of the KNCSP [14,15], presented results from selected institutions [15-17], or focused on specific cohorts [18]. Thus, this study aimed to confirm trends in the performance outcomes of mammography in the KNCSP over 12 years and investigate regional variations in performance in Korea.

MATERIALS AND METHODS

Data Source and Study Population

This study was approved by the Institutional Review Board of the College of Medicine, Catholic University of Korea (IRB No. MC21EESE0126). Informed consent was not required owing to the retrospective nature of the study.

The results of the KNCSP for BC were compiled in a database housed in the Korean National Health Insurance Big Data Base [15]. The KNCSP database includes mammography screening results; we analyzed BC screening data from 42165405 women, aged 40 years, who underwent screening via the KNCSP between 2009 and 2020.

Outcome Measurements

The mammographic screening results were classified into four categories according to the KNCSP guidelines for BC: negative (BI-RADS category 1), benign breast disease (BI-RADS category 2), suspected breast cancer (BI-RADS categories 4 and 5), and incomplete with additional evaluation (BI-RADS categories 0 and 3) [19]. We considered “negative” and “benign breast disease” as negative results, and “suspected breast cancer” and “incomplete” as positive results [15-19].

Following a screening mammogram, BC was identified based on the International Classification of Disease, 10th Revision codes C50 and D05, within one year after screening, based on the data from the co-payment reduction program [12]. The Korean government implemented this program to manage catastrophic medical costs and address unmet healthcare needs resulting from high expenditures on inpatient and outpatient services for individuals diagnosed with cancer, cerebrovascular disease, heart disease, rare diseases, severe incurable diseases, and severe burns [20,21]. Individuals diagnosed with a new cancer in Korea can benefit from reduced out-of-pocket payments for medical services through the National Health Insurance Program [21].

The performance metrics assessed included the RR, PPV, sensitivity, specificity, false positive rate (FPR), CDR, and interval cancer rate (ICR). The RR was calculated as the percentage of screened women who were recalled for further evaluation [17]. The PPV for the diagnosis of BC was defined as the percentage of mammographic examinations (BI-RADS categories 0, 3, 4, and 5) with positive findings that led to a pathological diagnosis of BC within 1 year. The CDR was calculated as the number of BC cases detected per 1000 examinations. ICR was defined as a histologically proven, invasive, or in situ cancer within one year following a negative screening test result and was calculated as the number of interval cancers per 1000 negative examinations [17]. Additionally, we reported breast density, as it influences mammographic interpretation [22]. The dense breast rate (DBR) was calculated using the 4th edition BI-RADS density categories, and those with BI-RADS categories C and D were considered to have dense breasts [13,19,23].

Korean Journal of Radiology KJR

Statistical Analysis

First, the performance results for each metric were analyzed from 2009 to 2020 using segmented regression (Joinpoint regression). We computed the annual percentage change (APC) and its 95% confidence interval (CI) and the average annual percent change (AAPC) from 2009 to 2020 [24]. The AAPC is based on the weighted average of the APC calculated across joinpoints. The Joinpoint model uses statistical criteria to determine when and how often the APC changes [25]. A linear model with a maximum of two joinpoints was used in the analysis. Each joinpoint denotes a statistically significant change in the trend [24].

Second, regional variations in mammography performance were analyzed across 237 districts (Si, Gun, and Gu), excluding 13 districts without BC screening institutions. A 'Si' is a small-to-medium-sized city outside of a metropolitan area, a 'Gun' is a district in less urbanized area, and a 'Gu' is a district within a large city [26]. Box plots were used to visualize the variation in performance by region over four years (2009, 2013, 2017, and 2020). All analyses were stratified into four age groups: 40–49, 50–59, 60–69, and 70 years.

The statistical significance level was set at $P < 0.05$. All statistical analyses were performed using the SAS statistical software (version 9.4; SAS Institute Inc., Cary, NC, USA). The crude rates obtained from SAS were input into the Joinpoint program (version 5.2.0.0. April 2024; Statistical Methodology and Applications Branch, Surveillance Research Program, National Cancer Institute, Bethesda, MD, USA), where APC and AAPC were calculated.

RESULTS

Mammography Utilization in a National Breast Cancer Screening Program, 2009–2020

Between 2009 and 2020, the number of participants who underwent mammography was 42165405, increasing from 2821132 in 2009 to 3596204 in 2020 (Table 1). Among women who underwent BC screening, 30.5% were 40–49 years, 32.2% were 50–59 years, 23.2% were 60–69 years, and 14.1% were ≥70 years. In 2009, women aged 40–49 represented 34.6% of BC screening participants, while those aged ≥70 comprised 11.6%. However, in 2020, the percentage of women aged 40–49 participating in screenings decreased to 26.5%, whereas that of women aged ≥70 increased to 15.3%.

Performance Results and DBR of Screening Mammography, 2009–2020

In 2020, the RR and PPV were 11.2% and 2.8%, respectively (Table 2). The sensitivity, specificity, and FPR were 68.9%, 89.1%, and 10.9%, respectively. The CDR and ICR were 3.1 per 1000 and 1.6 per 1000, respectively. The DBR was 55.9%.

The trend in performance metric results for mammographic screening at the KNCSP over 12 years was as follows (Fig. 1): the RR showed a downward trend after 2017 (APC [95% CI], -8.0% [-10.8% to -5.8%]). In contrast, an overall upward trend was observed in the PPV, which further increased after 2017 (APC: 9.2% from 2009 to 2017 and 14.8% from 2017 to 2020). Sensitivity did not change significantly after 2014,

Table 1. Number of screening institutions and participants undergoing mammography by age and year

Screening year	Institutions	Participants by age group				
		Total	40–49 years	50–59 years	60–69 years	70 years or over
2009	1954	2821132	976790 (34.6)	910673 (32.3)	606706 (21.5)	326963 (11.6)
2010	2128	2758047	949646 (34.4)	891596 (32.3)	581643 (21.1)	335162 (12.2)
2011	2288	3228705	1042572 (32.3)	1104847 (34.2)	668966 (20.7)	412320 (12.8)
2012	2404	3214715	1010201 (31.4)	1071681 (33.3)	690122 (21.5)	442711 (13.8)
2013	2467	3290138	1046324 (31.8)	1093482 (33.2)	692755 (21.1)	457577 (13.9)
2014	2501	3442834	1058931 (30.8)	1118436 (32.5)	767838 (22.3)	497629 (14.5)
2015	2546	3652977	1128600 (30.9)	1176299 (32.2)	834766 (22.9)	513312 (14.1)
2016	2624	3901778	1180533 (30.3)	1240360 (31.8)	918030 (23.5)	562855 (14.4)
2017	2736	3997314	1194269 (29.9)	1262357 (31.6)	949570 (23.8)	591118 (14.8)
2018	2850	4025231	1133736 (28.2)	1279159 (31.8)	999110 (24.8)	613226 (15.2)
2019	2941	4236330	1205925 (28.5)	1298691 (30.7)	1069756 (25.3)	661958 (15.6)
2020	3010	3596204	952259 (26.5)	1108906 (30.8)	983903 (27.4)	551136 (15.3)
Total	30449	42165405	12879786 (30.5)	13556487 (32.2)	9763165 (23.2)	5965967 (14.1)

Values are presented as number (%)

Table 2. Performance metrics of breast cancer screening and DBR by year

Indicators	2009	2010	2011	2012	2013	2014	2015	2016	2017	2018	2019	2020	AAPC (95% CI)
RR, %	17.18	15.05	14.56	14.66	14.53	14.91	14.61	14.17	13.99	13.02	11.92	11.17	-3.73% (-4.21, -3.2)*
PPV, %	0.84	1.00	1.06	1.19	1.30	1.36	1.50	1.64	1.80	2.05	2.30	2.77	10.7% (9.66, 11.41)*
Sensitivity, %	67.30	65.32	64.39	66.09	67.15	68.34	68.68	67.64	68.81	68.86	68.89	68.93	0.25% (0.12, 0.39)*
Specificity, %	82.93	85.07	85.56	85.48	85.62	85.25	85.56	86.01	86.21	87.20	88.31	89.09	0.62% (0.52, 0.71)*
FPR, %	17.07	14.93	14.44	14.52	14.38	14.75	14.44	13.99	13.79	12.80	11.69	10.91	-3.87% (-4.33, -3.35)*
CDR, ‰	1.45	1.50	1.54	1.75	1.89	2.02	2.19	2.33	2.51	2.67	2.75	3.10	7.30% (6.79, 7.81)*
ICR, ‰	0.85	0.94	1.00	1.05	1.08	1.10	1.17	1.30	1.32	1.39	1.41	1.57	5.19% (4.48, 5.91)*
DBR, %	40.50	44.72	45.93	47.40	48.77	49.76	50.74	52.09	53.73	54.16	55.16	55.89	2.96% (2.68, 3.19)*

*$P < 0.05$.
DBR = dense breast rate, AAPC = average annual percent change, CI = confidence interval, RR = recall rate, PPV = positive predictive value, FPR = false positive rate, CDR = cancer detection rate per 1000 examinations, ICR = interval cancer rate per 1000 negative examinations

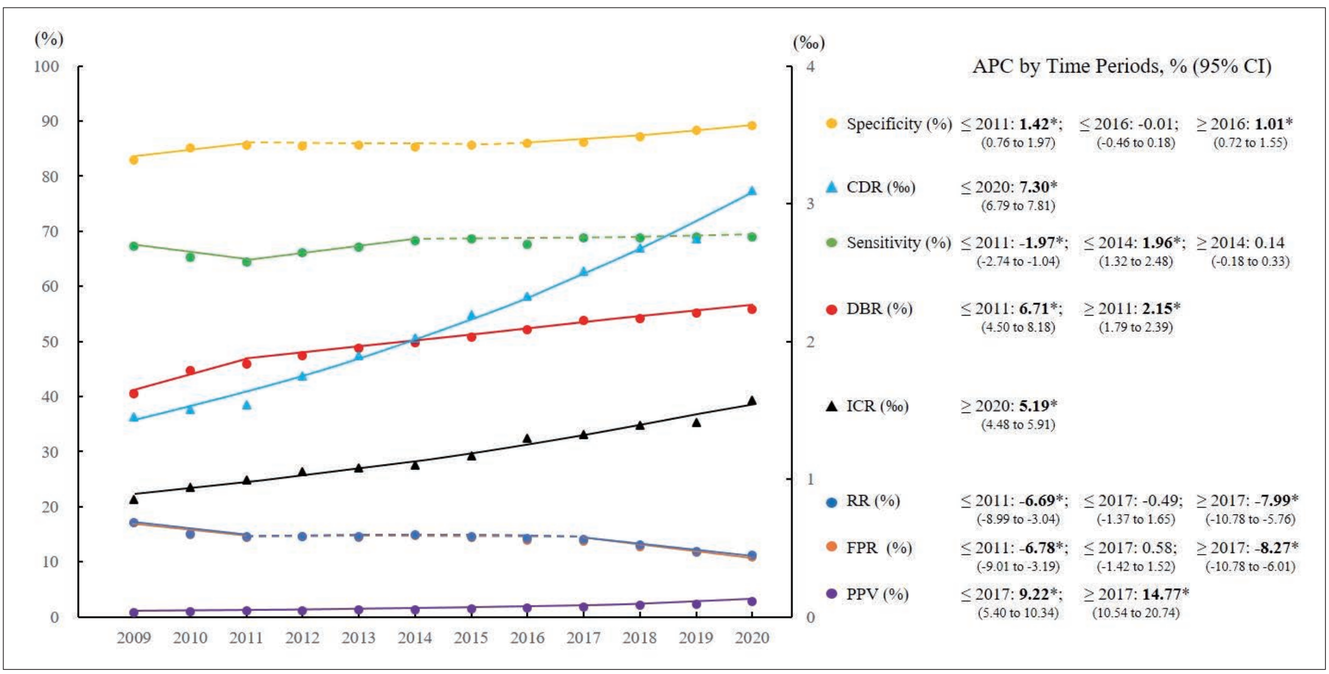


Fig. 1. Trends in breast cancer screening performance and DBR, 2009–2020. The APC provided on the right side of the figure, with values expressed as percentages, was calculated to quantitatively measure longitudinal changes from baseline to the joinpoint. A linear model with a maximum of two joinpoints was used in the analysis. A negative APC value indicated a decrease over time, whereas a positive value indicated an increase. *$P < 0.05$, corresponds to the periods indicated by solid lines in the graphs; the periods marked with dotted lines are statistically insignificant. DBR = dense breast rate, APC = annual percentage change, CI = confidence interval, CDR = cancer detection rate per 1000 examinations, ICR = interval cancer rate per 1000 negative examinations, RR = recall rate, FPR = false-positive rate, PPV = positive predictive value

specificity increased slightly after 2016 (APC, 1.0% from 2016 to 2020), and the FPR showed downward trends after 2017 (APC [95% CI], -8.3% [-10.8% to -6.0%]). The CDR and ICR showed continuous upward trends after 2009 (APC [95% CI]: 7.3% [6.8% to 7.8%] and 5.2% [4.5% to 5.9%], respectively). Finally, although the DBR slowed after 2011, a general upward trend was observed (APC: 6.7% from 2009 to 2011 and 2.2% from 2011 to 2020).

Age-Stratified Performance Results of Screening Mammography, 2009–2020

Figures 2–5 and Supplementary Tables 1–4 shows the trends in BC screening performance according to age group. The RR, FPR, CDR, ICR, and DBR were higher in women aged 40–49 years than in older women, whereas the PPV, sensitivity, and specificity were comparatively lower. Over time, the PPV, DBR, CDR, and ICR increased across all age groups. The metrics that demonstrated age-related

https://doi.org/10.3348/kjr.2024.0866

Korean Journal of Radiology
KJR

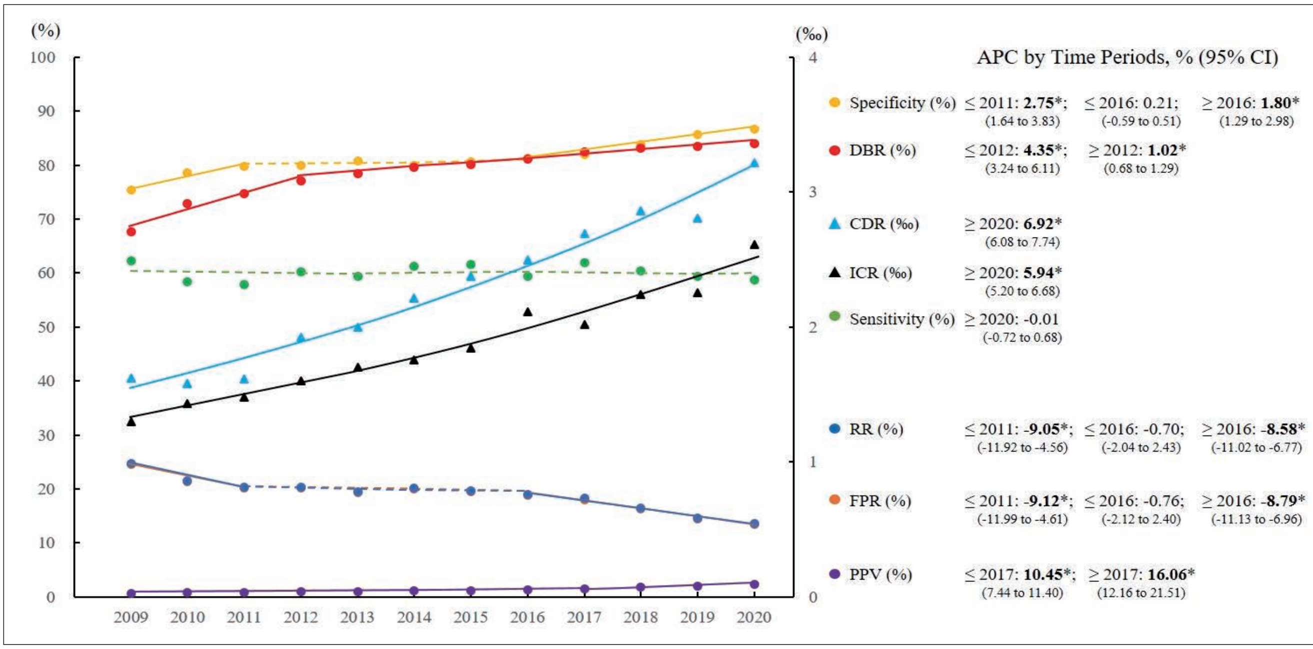


Fig. 2. Age-stratified performance in breast cancer screening for women aged 40–49 years, 2009–2020. The APC provided on the right side of the figure, with values expressed as percentages, was calculated to quantitatively measure longitudinal changes from baseline to the joinpoint. A linear model with a maximum of two joinpoints was used in the analysis. A negative APC value indicated a decrease over time, whereas a positive value indicated an increase. *$P < 0.05$, corresponds to the periods indicated by solid lines in the graphs; the periods marked with dotted lines are statistically insignificant. APC = annual percentage change, CI = confidence interval, DBR = dense breast rate, CDR = cancer detection rate per 1000 examinations, ICR = interval cancer rate per 1000 negative examinations, RR = recall rate, FPR = false-positive rate, PPV = positive predictive value

differences in the trends were the RR, FPR, specificity, and sensitivity. The RR and FPR decreased in women aged 40–49 years, after 2016, and in women aged ≥50 years, after 2017. The specificity increased in women aged 40–49 years after 2016, 50–59 years after 2015, and 60 years after 2017. Sensitivity showed no statistically significant difference between women in their 40s and 50s; however, it steadily increased among women over 60 years in the past 12 years.

Regional Variations in Performance Metrics and DBRs of Screening Mammography

Figure 6 illustrates variations in performance metrics across 237 regions in Korea for 2009, 2013, 2017, and 2020. The standard deviation (SD) and interquartile range (IQR) of the RR for mammography screening in 2009 were 11.1 and 14.3, respectively, which decreased to 6.4 and 6.8 by 2020. The SD of the PPV increased from 1.9 in 2009 to 6.8 in 2020, and the IQR increased from 0.8–2.4. The SD and IQR of sensitivity, specificity, and FPR decreased in 2020 compared with 2009, whereas the SD and IQR of the CDR and ICR increased in 2020 compared to 2009.

We also analyzed the distribution of performance indicators stratified by age group (Supplementary Table 5). Women aged 40–49 years exhibited greater regional variations in the RR, specificity, FPR, and ICR than those observed in older age groups. Women aged ≥70 years had an increased IQR of RR, specificity, FPR, and DBR compared to women in younger age groups.

DISCUSSION

We analyzed 42 million screening mammograms performed on women aged ≥40 years in Korea through the KNCSP from 2009 to 2020. First, we compared the performance of the mammography screening program in Korea with international benchmarks and examined trends in performance metrics over 12 years. In 2020, mammography performance metrics showed significant improvements compared to 2009. The RR (11.2%), specificity (89.1%), and CDR (3.1‰) were within acceptable ranges for the NMD and met BCSC [6-9] benchmarks. However, the sensitivity (68.9%) and PPV (2.8%) remained low [6-9]. Second, we analyzed regional variations in mammography performance across Korea's districts in four years: 2009, 2013, 2017, and

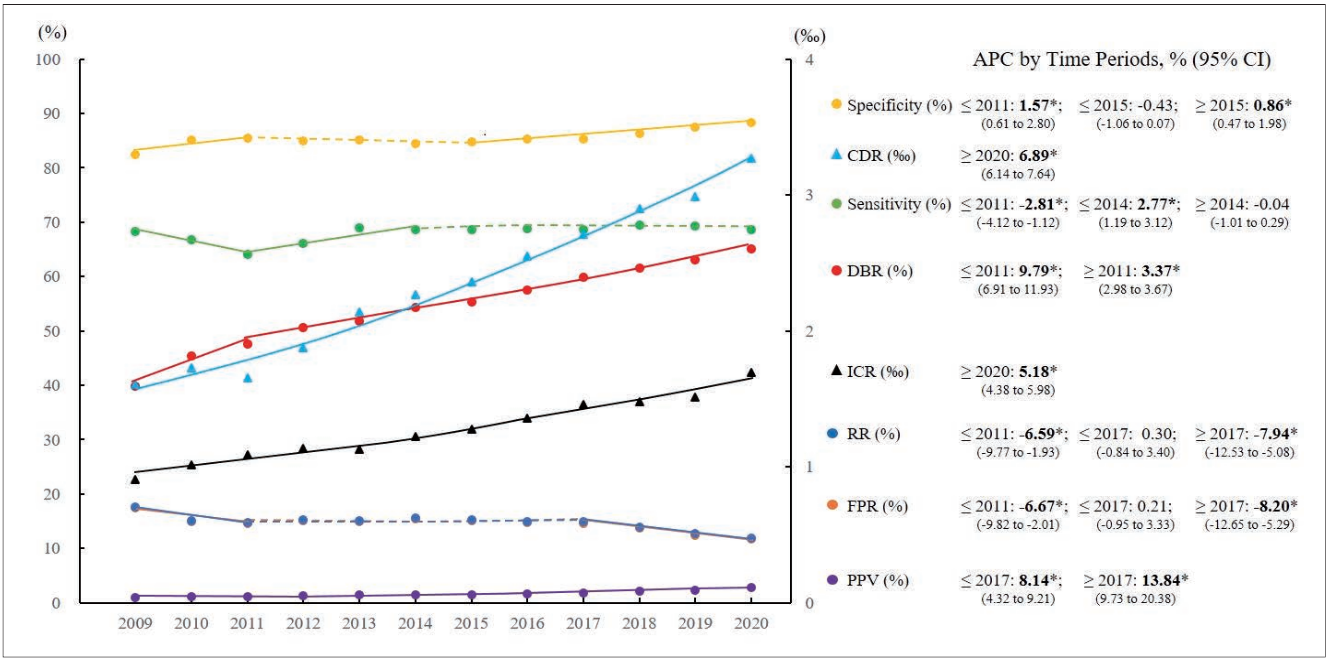


Fig. 3. Age-stratified performance in breast cancer screening for women aged 50–59 years, 2009–2020. The APC provided on the right side of the figure, with values expressed as percentages, was calculated to quantitatively measure longitudinal changes from baseline to the joinpoint. A linear model with a maximum of two joinpoints was used in the analysis. A negative APC value indicated a decrease over time, whereas a positive value indicated an increase. **P* < 0.05, corresponds to the periods indicated by solid lines in the graphs; the periods marked with dotted lines are statistically insignificant. APC = annual percentage change, CI = confidence interval, CDR = cancer detection rate per 1000 examinations, DBR = dense breast rate, ICR = interval cancer rate per 1000 negative examinations, RR = recall rate, FPR = false-positive rate, PPV = positive predictive value

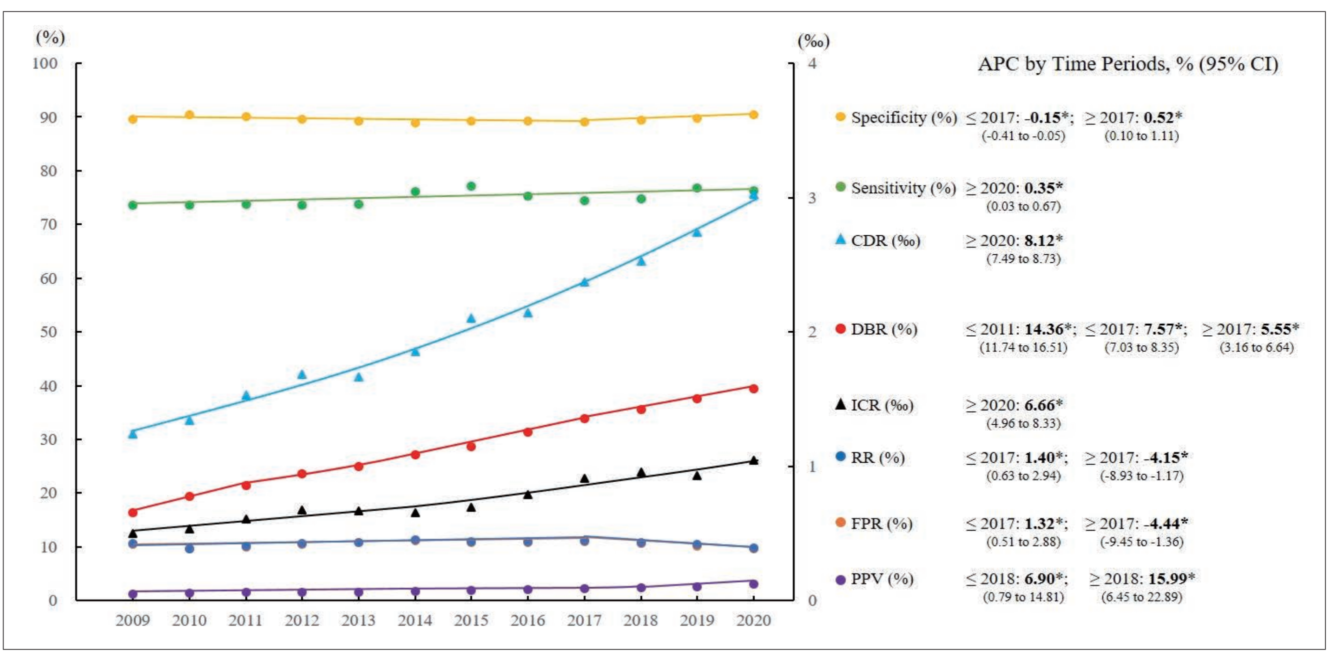


Fig. 4. Age-stratified performance in breast cancer screening for women aged 60–69 years, 2009–2020. The APC provided on the right side of the figure, with values expressed as percentages, was calculated to quantitatively measure longitudinal changes from baseline to the joinpoint. A linear model with a maximum of two joinpoints was used in the analysis. A negative APC value indicated a decrease over time, whereas a positive value indicated an increase. **P* < 0.05, corresponds to the periods indicated by solid lines in the graphs; the periods marked with dotted lines are statistically insignificant. APC = annual percentage change, CI = confidence interval, CDR = cancer detection rate per 1000 examinations, DBR = dense breast rate, ICR = interval cancer rate per 1000 negative examinations, RR = recall rate, FPR = false-positive rate, PPV = positive predictive value

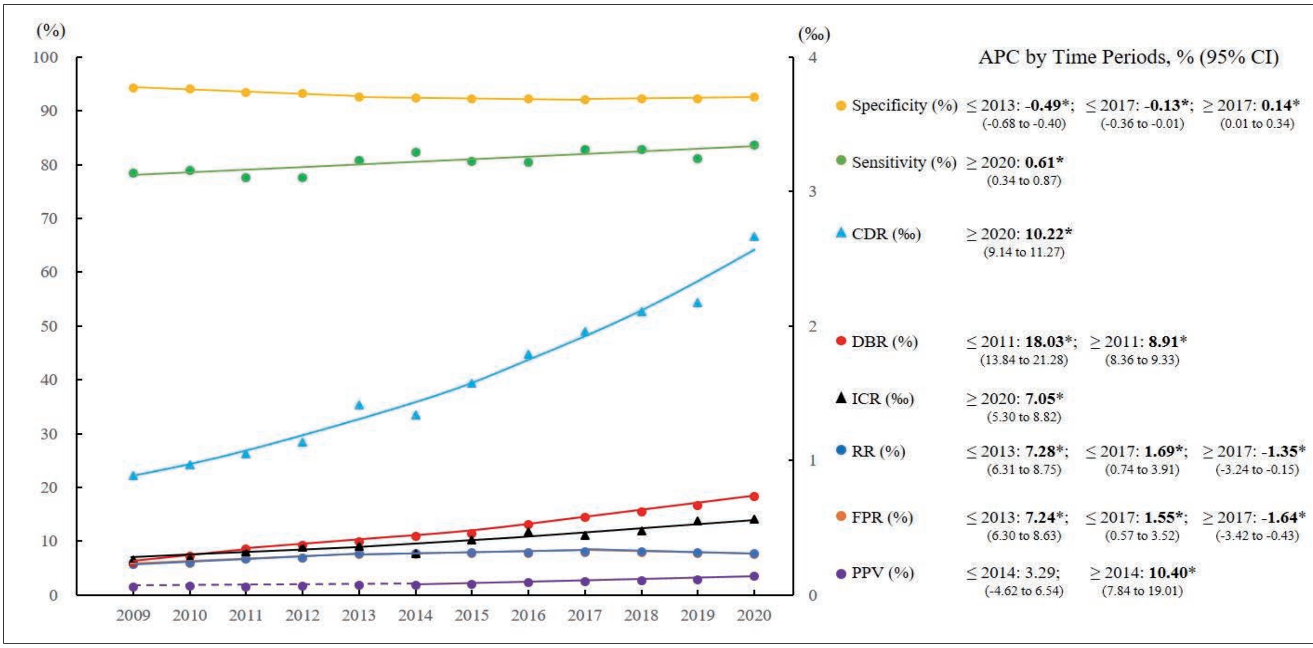


Fig. 5. Age-stratified performance in breast cancer screening for women aged 70 or older, 2009–2020. The APC provided on the right side of the figure, with values expressed as percentages, was calculated to quantitatively measure longitudinal changes from baseline to the joinpoint. A linear model with a maximum of two joinpoints was used in the analysis. A negative APC value indicated a decrease over time, whereas a positive value indicated an increase. *$P < 0.05$, corresponds to the periods indicated by solid lines in the graphs; the periods marked with dotted lines are statistically insignificant. APC = annual percentage change, CI = confidence interval, CDR = cancer detection rate per 1000 examinations, DBR = dense breast rate, ICR = interval cancer rate per 1000 negative examinations, RR = recall rate, FPR = false-positive rate, PPV = positive predictive value

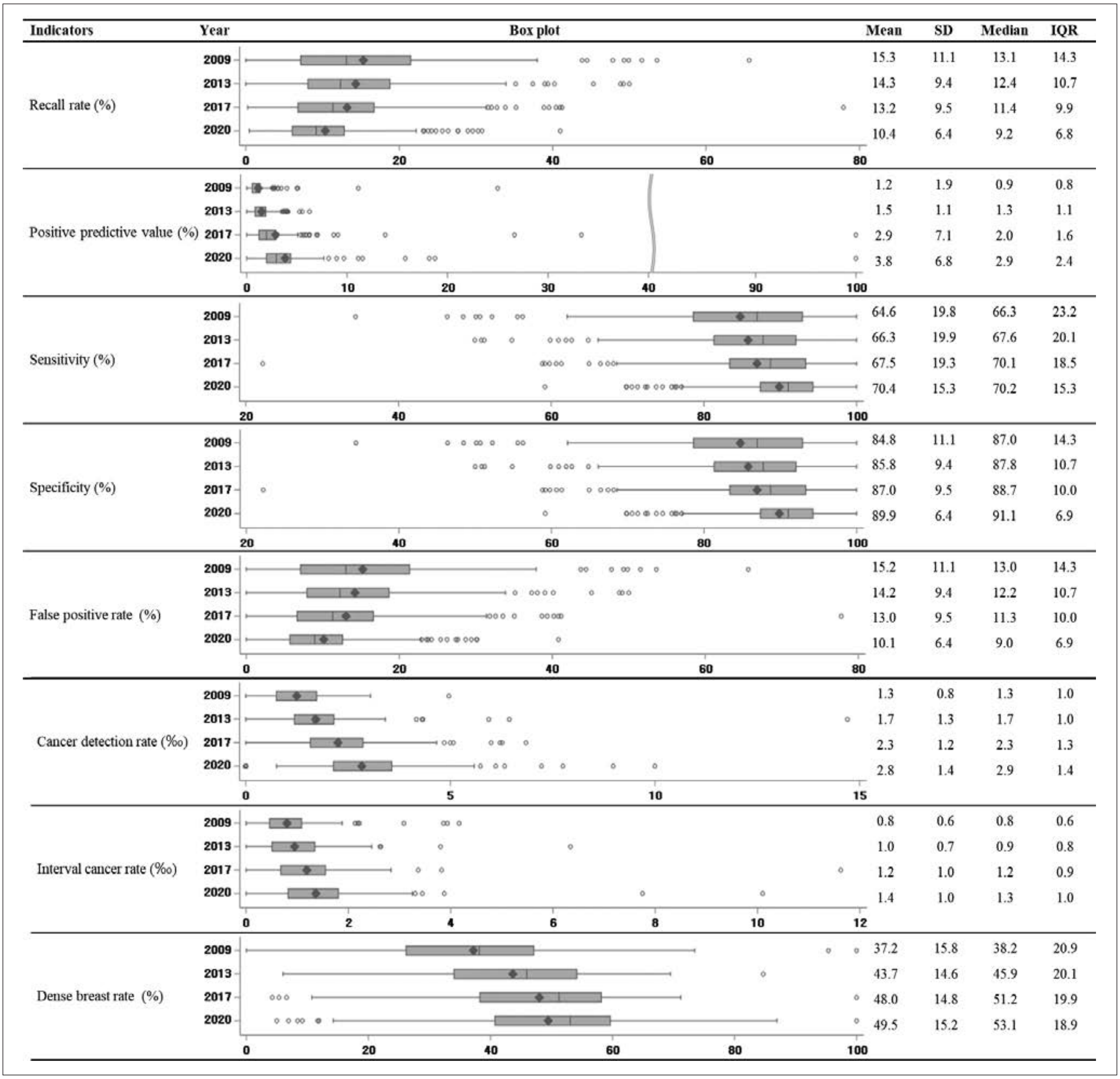
2020. Over time, regional differences in the RR, sensitivity, specificity, and FPR decreased, whereas variations in the PPV, CDR, and ICR increased.

Among the performance metrics for mammography reported by Lee et al. [15] in Korea, the sensitivity in 2010 was 80.2%, whereas our results showed a sensitivity of 65.3%. The previous study differed in that it was based on 130000 mammography results (2005–2010) from 11 institutions [15]. In contrast, our sensitivity results are comparable to those observed in a cohort from a screening center (2011–2018) with 200000 cases, which reported a sensitivity of 64.0% for women in their 40s and a sensitivity of 79.8% for women aged ≥50 years [18]. Compared to Japan, another East Asian country, in 2009, the RR (10.5% vs. 17.2%) was higher, while the sensitivity (73.7% vs. 67.3%), specificity (89.4% vs. 82.9%), CDR (2.8% vs. 1.5%), and PPV (2.6% vs. 0.8%) were relatively lower [27].

In BC screening studies, cancer is considered an interval cancer if it is diagnosed within a defined time period after screening mammography (e.g., 1, 2, or 3 years), depending on the study. In Canada, annual screening for women with dense breasts resulted in an ICR of 0.9 per 1000 compared with an incidence rate of 2.9 per 1000 for women following standard biennial screening policies [28]. The ICR in the UK was 3.1 per 1000 women at 3-year intervals [29]. The ICR for mammography, including various screening intervals, reportedly ranges from 0.5–5 per 1000. The ICR in Korea has steadily increased since 2009, reaching 1.6 in 2020 [28-30]. To maintain continuity and comparability with previous studies conducted in Korea, we defined the period of interval cancer occurrence as one year [15-17]. Screening techniques and intervals can influence the ICR [30,31]. The ICR can vary as much as 28% when the most and least restrictive definitions are used [32]. Therefore, caution should be exercised when comparing the ICR of other BC screening programs, and various intervals should be calculated collectively [30-32].

In large-scale screening mammography programs, reducing recalls can increase the ICR and PPV because the probability of false-positive results increases with the number of recalls [29,33,34]. Over the past 12 years, the RR has decreased, whereas the PPV and ICR have increased in Korea. These changes have been significant since 2017. Previous literature reviews have shown that increases in the

 https://doi.org/10.3348/kjr.2024.0866



Indicators	Year	Mean	SD	Median	IQR
Recall rate (%)	2009	15.3	11.1	13.1	14.3
	2013	14.3	9.4	12.4	10.7
	2017	13.2	9.5	11.4	9.9
	2020	10.4	6.4	9.2	6.8
Positive predictive value (%)	2009	1.2	1.9	0.9	0.8
	2013	1.5	1.1	1.3	1.1
	2017	2.9	7.1	2.0	1.6
	2020	3.8	6.8	2.9	2.4
Sensitivity (%)	2009	64.6	19.8	66.3	23.2
	2013	66.3	19.9	67.6	20.1
	2017	67.5	19.3	70.1	18.5
	2020	70.4	15.3	70.2	15.3
Specificity (%)	2009	84.8	11.1	87.0	14.3
	2013	85.8	9.4	87.8	10.7
	2017	87.0	9.5	88.7	10.0
	2020	89.9	6.4	91.1	6.9
False positive rate (%)	2009	15.2	11.1	13.0	14.3
	2013	14.2	9.4	12.2	10.7
	2017	13.0	9.5	11.3	10.0
	2020	10.1	6.4	9.0	6.9
Cancer detection rate (‰)	2009	1.3	0.8	1.3	1.0
	2013	1.7	1.3	1.7	1.0
	2017	2.3	1.2	2.3	1.3
	2020	2.8	1.4	2.9	1.4
Interval cancer rate (‰)	2009	0.8	0.6	0.8	0.6
	2013	1.0	0.7	0.9	0.8
	2017	1.2	1.0	1.2	0.9
	2020	1.4	1.0	1.3	1.0
Dense breast rate (%)	2009	37.2	15.8	38.2	20.9
	2013	43.7	14.6	45.9	20.1
	2017	48.0	14.8	51.2	19.9
	2020	49.5	15.2	53.1	18.9

Fig. 6. Regional variations in performance of screening mammography. The bottom, middle, and top borders of the boxes represent the 25th, 50th, and 75th percentiles across all districts, while circles indicate outliers and diamonds represent national averages. SD = standard deviation, IQR = interquartile range

RR are not directly associated with an improved CDR [33,35]. However, the CDR may increase as the RR increases within a low range of RR (up to approximately 8%) [33,36]. Korea's RR has steadily declined but remains higher than other benchmarks [6-9]. Additionally, cancer incidence affects performance metrics, including the CDR and ICR [37]. The number of women newly diagnosed with BC in Korea has increased from 13399 in 2009 to 24806 in 2020 [38]. The performance metrics of mammography should be interpreted considering the increasing incidence of BC.

Our study showed that the proportion of dense breasts increased steadily from 2009 to 2020. An increase in the DBR among women is associated with changes in the prevalence of several factors, including smoking, drinking, lack of exercise, early menarche, premenopausal status, nulliparity, no history of breastfeeding, and obesity [39]. These factors associated with dense breasts and BC incidence overlapped, and 96% of the variation in BC incidence was explained by the variation in the prevalence of dense breasts in Korea [39]. Additionally, more than half of the women in our study were in their 40s and 50s and the proportion of dense breasts was higher than that of fatty breasts. Younger age and higher breast density were associated with higher RRs, FPRs, and ICRs and lower PPVs and specificities, negatively affecting performance in mammography screening [16,33,39]. Compared to other age groups, women in their 40s exhibited the most notable decrease in RRs and the most significant increase in PPVs, while the CDR and ICR increased the most in those aged over 70 years. Although the CDR and ICR increase with breast density [16,40], the trend shift observed in women over 70 years of age, who generally exhibit lower breast density, is concerning [11]. Therefore, it is necessary to develop a tailored screening strategy that considers differences in performance metrics for BC screening based on age.

Regional variation in healthcare serve is an indirect measure of performance and quality within the healthcare system [41,42]. In Western countries, disparities in mammography screening utilization have been reported at the city level [43-45]. Although an earlier study could not explain the source of the association between screening mammography performance and geography, it identified variations in the geographic location of radiology practices, suggesting the need for further investigation [6]. Consequently, investigating geographic variations in mammography screening may help improve early detection strategies [45,46].

We observed variations in mammography performance metrics across administrative district in Korea. The regional variation in the RR, sensitivity, specificity, and FPR for mammography screening decreased from 2009 to 2020. In contrast, the regional variation in the PPV, CDR, and ICR increased during the same period [47], exhibiting an annual upward trend at low mean values. When mean values are low, even small absolute changes can lead to large relative changes, making regional variations more volatile, particularly in metrics with low mean values. Our results could serve as baseline data for establishing regional targets to enhance the quality of mammography screening.

Our study has some limitations. First, we did not include mammography in opportunistic cancer screening. In 2019, 63.8% of the eligible population participated in the KNCSP to assess the presence of BC [48]. Second, we could not distinguish between pathological types of invasive breast carcinoma. The most prevalent type of BC is categorized into invasive ductal carcinoma (IDC) and ductal carcinoma in situ (DCIS). Pure IDC and IDC combined with DCIS may exhibit different biological behaviors, potentially leading to varying outcomes [49]. Third, the performance of mammography may be related to the frequency of national screening mammography between 2009 and 2020. However, we did not account for the variability in participants' screening frequency. Fourth, due to limited resources, we did not evaluate associations between equipment type or participant factors that could influence mammography. Further studies involving larger populations are needed to determine differences in mammography performance based on participant factors and equipment types.

In conclusion, this research provides crucial baseline data for BC screening management. The specificity and CDR for screening mammography in the KNCSP from 2009 to 2020 were comparable to those observed in Western women. However, the RR was suboptimal despite showing improvements over this period. The PPV and sensitivity remained low, indicating the need for continuous monitoring of these metrics in the future. Furthermore, although regional variations in RR, sensitivity, specificity, and FPR decreased in 2020 compared to 2009, areas that have consistently been outliers need to be identified and addressed. Additional analyses are also necessary to elucidate the factors contributing to these variations.

Supplement

The Supplement is available with this article at https://doi.org/10.3348/kjr.2024.0866.

Availability of Data and Material

The datasets generated or analyzed during the study are based on the NHIS database, so that they are available from NHIS on reasonable request.

Conflicts of Interest

The authors have no potential conflicts of interest to disclose.

Author Contributions

Conceptualization: Dong Jun Kim, Nan-He Yoon, Jae Kwan Jun, Mina Suh, Hooyeon Lee. Data curation: Dong Jun Kim, Sunhwa Lee, Seongju Kim, Horim A. Hwang. Formal analysis: Dong Jun Kim, Nan-He Yoon, Sunhwa Lee. Funding acquisition: Hooyeon Lee. Investigation: Dong Jun Kim, Seongju Kim, Sunhwa Lee. Methodology: Dong Jun Kim, Nan-He Yoon, Jae Kwan Jun, Mina Suh, Hooyeon Lee. Project administration: Dong Jun Kim, Sunhwa Lee, Seongju Kim, Hooyeon Lee. Resources: Dong Jun Kim, Seongju Kim, Horim A. Hwang, Hooyeon Lee. Software: Dong Jun Kim, Sunhwa Lee, Sunhwa Lee, Seongju Kim. Supervision: Jae Kwan Jun, Mina Suh, Seung Eun Jung, Hooyeon Lee. Validation: all authors. Visualization: Dong Jun Kim, Seongju Kim, Horim A. Hwang. Writing—original draft: Dong Jun Kim, Nan-He Yoon, Hooyeon Lee. Writing—review & editing: all authors.

ORCID IDs
Dong Jun Kim
https://orcid.org/0000-0002-8015-9948
Nan-He Yoon
https://orcid.org/0000-0003-1268-0149
Jae Kwan Jun
https://orcid.org/0000-0003-1647-0675
Mina Suh
https://orcid.org/0000-0001-8101-7493
Sunhwa Lee
https://orcid.org/0009-0007-9269-8344
Seongju Kim
https://orcid.org/0000-0002-0519-1278
Horim A. Hwang
https://orcid.org/0000-0002-3616-0810
Seung Eun Jung
https://orcid.org/0000-0003-0674-5444
Hooyeon Lee
https://orcid.org/0000-0002-8426-2045

Funding Statement
This study was supported by the National R&D Program for Cancer Control through the National Cancer Center (NCC) funded by the Ministry of Health & Welfare, Republic of Korea (HA21C0225).

REFERENCES

1. Arnold M, Morgan E, Rumgay H, Mafra A, Singh D, Laversanne M, et al. Current and future burden of breast cancer: global statistics for 2020 and 2040. *Breast* 2022;66:15-23
2. Grunfeld E, Noorani H, McGahan L, Paszat L, Coyle D, van Walraven C, et al. Surveillance mammography after treatment of primary breast cancer: a systematic review. *Breast* 2002;11:228-235
3. Monticciolo DL, Malak SF, Friedewald SM, Eby PR, Newell MS, Moy L, et al. Breast cancer screening recommendations inclusive of all women at average risk: update from the ACR and Society of Breast Imaging. *J Am Coll Radiol* 2021;18:1280-1288
4. Marmot MG, Altman DG, Cameron DA, Dewar JA, Thompson SG, Wilcox M. The benefits and harms of breast cancer screening: an independent review. *Br J Cancer* 2013;108:2205-2240
5. D'Orsi CJ, Sickles EA, Mendelson EB, Morris EA. *2013 ACR BI-RADS atlas: breast imaging reporting and data system*. Reston: American College of Radiology, 2014
6. Lee CS, Moy L, Hughes D, Golden D, Bhargavan-Chatfield M, Hemingway J, et al. Radiologist characteristics associated with interpretive performance of screening mammography: a national mammography database (NMD) study. *Radiology* 2021;300:518-528
7. Lee CS, Bhargavan-Chatfield M, Burnside ES, Nagy P, Sickles EA. The national mammography database: preliminary data. *AJR Am J Roentgenol* 2016;206:883-890
8. Miglioretti DL, Ichikawa L, Smith RA, Bassett LW, Feig SA, Monsees B, et al. Criteria for identifying radiologists with acceptable screening mammography interpretive performance on basis of multiple performance measures. *AJR Am J Roentgenol* 2015;204:W486-W491
9. Lehman CD, Arao RF, Sprague BL, Lee JM, Buist DS, Kerlikowske K, et al. National performance benchmarks for modern screening digital mammography: update from the breast cancer surveillance consortium. *Radiology* 2017;283:49-58
10. Yap YS, Lu YS, Tamura K, Lee JE, Ko EY, Park YH, et al. Insights into breast cancer in the east vs the west: a review. *JAMA Oncol* 2019;5:1489-1496
11. Armstrong K, Eisen A, Weber B. Assessing the risk of breast cancer. *N Engl J Med* 2000;342:564-571
12. Hong S, Song SY, Park B, Suh M, Choi KS, Jung SE, et al. Effect of digital mammography for breast cancer screening: a comparative study of more than 8 million Korean women. *Radiology* 2020;294:247-255
13. Ministry of Health and Welfare, National Cancer Center. *Quality guidelines of breast cancer screening*. 2nd ed. Seoul: Ministry of Health and Welfare, National Cancer Center, 2018
14. Choi N, Lee A, Lee HK, Yi BH, Cha JG. [Mammogrphic screening provided by the national health insurance corporation: a 1 year audit in a secondary medical institution]. *J Korean Soc Radiol* 2009;60:51-55. Korean
15. 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
16. Kim YJ, Lee EH, Jun JK, Shin DR, Park YM, Kim HW, et al. 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
17. Choi BH, Lee EH, Jun JK, Kim KW, Park YM, Kim HW, et al. Effect of different types of mammography equipment on screening outcomes: a report by the alliance for breast cancer screening in Korea. *Korean J Radiol* 2019;20:1638-1645
18. Kwon MR, Chang Y, Park B, Ryu S, Kook SH. Performance

analysis of screening mammography in Asian women under 40 years. *Breast Cancer* 2023;30:241-248
19. Kim S, Tran TXM, Song H, Ryu S, Chang Y, Park B. Mammographic breast density, benign breast disease, and subsequent breast cancer risk in 3.9 million Korean women. *Radiology* 2022;304:534-541
20. Shin DW, Cho J, Park JH, Cho B. National general health screening program in Korea: history, current status, and future direction. *Precis Future Med* 2022;6:9-31
21. Kim DJ, Yoon NH, Jun JK, Suh M, Lee S, Kim S, et al. Association between endoscopist volume and interval cancers after colonoscopy: results from the national colorectal cancer screening program in Korea. *Cancer Res Treat* 2024;56:1164-1170
22. Youk JH, Kim SJ, Son EJ, Gweon HM, Kim JA. Comparison of visual assessment of breast density in BI-RADS 4th and 5th editions with automated volumetric measurement. *AJR Am J Roentgenol* 2017;209:703-708
23. D'Orsi C, Bassett L, Feig S. *Breast imaging reporting and data system (BI-RADS)*. 4th ed. Reston: American College of Radiology, 2018
24. Kim HJ, Fay MP, Feuer EJ, Midthune DN. Permutation tests for joinpoint regression with applications to cancer rates. *Stat Med* 2000;19:335-351
25. Kim DJ, Cho KH, Kim S, Lee H. Influenza vaccination before and during the COVID-19 pandemic in the elderly in South Korea. *J Infect Public Health* 2024;17:503-508
26. Choi K. Job creation during Korea's transition to a knowledge economy. *KDI J Econ Policy* 2022;44:75-99
27. Ohta K, Kasahara Y, Tanaka F, Maeda H. Is clinical breast examination effective in Japan? Consideration from the age-specific performance of breast cancer screening combining mammography with clinical breast examination. *Breast Cancer* 2016;23:183-189
28. Seely JM, Peddle SE, Yang H, Chiarelli AM, McCallum M, Narasimhan G, et al. Breast density and risk of interval cancers: the effect of annual versus biennial screening mammography policies in Canada. *Can Assoc Radiol J* 2022;73:90-100
29. Burnside ES, Vulkan D, Blanks RG, Duffy SW. Association between screening mammography recall rate and interval cancers in the UK breast cancer service screening program: a cohort study. *Radiology* 2018;288:47-54
30. Zeng A, Brennan ME, Young S, Mathieu E, Houssami N. The effect of supplemental imaging on interval cancer rates in mammography screening: systematic review. *Clin Breast Cancer* 2022;22:212-222
31. Ray KM. Interval cancers in understanding screening outcomes. *Radiol Clin North Am* 2024;62:559-569
32. Loy EY, Molinar D, Chow KY, Fock C. National breast cancer screening programme, Singapore: evaluation of participation and performance indicators. *J Med Screen* 2015;22:194-200
33. Yankaskas BC, Klabunde CN, Ancelle-Park R, Renner G, Wang H, Fracheboud J, et al. International comparison of performance measures for screening mammography: can it be done? *J Med Screen* 2004;11:187-193
34. Otten JD, Karssemeijer N, Hendriks JH, Groenewoud JH, Fracheboud J, Verbeek AL, et al. Effect of recall rate on earlier screen detection of breast cancers based on the Dutch performance indicators. *J Natl Cancer Inst* 2005;97:748-754
35. Elmore JG, Nakano CY, Koepsell TD, Desnick LM, D'Orsi CJ, Ransohoff DF. International variation in screening mammography interpretations in community-based programs. *J Natl Cancer Inst* 2003;95:1384-1393
36. Yankaskas BC, Cleveland RJ, Schell MJ, Kozar R. Association of recall rates with sensitivity and positive predictive values of screening mammography. *AJR Am J Roentgenol* 2001;177:543-549
37. Selby K, Sedki M, Levine E, Kamineni A, Green BB, Vachani A, et al. Test performance metrics for breast, cervical, colon, and lung cancer screening: a systematic review. *J Natl Cancer Inst* 2023;115:375-384
38. Kang MJ, Jung KW, Bang SH, Choi SH, Park EH, Yun EH, et al. Cancer statistics in Korea: incidence, mortality, survival, and prevalence in 2020. *Cancer Res Treat* 2023;55:385-399
39. Kim S, Tran TXM, Park B. Trends in breast density and other risk factors for breast cancer and associations with trends in the incidence of breast cancer in Korean women. *Maturitas* 2024;189:108070
40. 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
41. Srivastava D, Lafortune G, Paris V, Belloni A. Geographic variations in health care [accessed on January 6, 2025]. Available at: https://doi.org/10.1787/9789264216594-en
42. Brownlee S, Chalkidou K, Doust J, Elshaug AG, Glasziou P, Heath I, et al. Evidence for overuse of medical services around the world. *Lancet* 2017;390:156-168
43. Kim E, Moy L, Gao Y, Hartwell CA, Babb JS, Heller SL. City patterns of screening mammography uptake and disparity across the United States. *Radiology* 2019;293:151-157
44. Ponti A, Anttila A, Ronco G, Senore C. Cancer screening in the European Union (2017) [accessed on January 6, 2025]. Available at: https://screening.iarc.fr/EUreport.php
45. Czwikla J, Urbschat I, Kieschke J, Schüssler F, Langner I, Hoffmann F. Assessing and explaining geographic variations in mammography screening participation and breast cancer incidence. *Front Oncol* 2019;9:909
46. Mobley LR, Kuo TM, Urato M, Subramanian S, Watson L, Anselin L. Spatial heterogeneity in cancer control planning and cancer screening behavior. *Ann Assoc Am Geogr* 2012;102:1113-1124
47. Citrome L. Relative vs. absolute measures of benefit and risk: what's the difference? *Acta Psychiatr Scand* 2010;121:94-102
48. Park H, Seo SH, Park JH, Yoo SH, Keam B, Shin A. The impact of COVID-19 on screening for colorectal, gastric, breast, and cervical cancer in Korea. *Epidemiol Health* 2022;44:e2022053
49. Kole AJ, Park HS, Johnson SB, Kelly JR, Moran MS, Patel AA. Overall survival is improved when DCIS accompanies invasive breast cancer. *Sci Rep* 2019;9:9934