Radiology

Reduction in False-Positive Results after Introduction of Digital Mammography: Analysis from Four Population-based Breast Cancer Screening Programs in Spain[1]

Maria Sala, MD, PhD
Dolores Salas, MD
Francesc Belvis, BSc
Mar Sánchez, MD
Joana Ferrer, MD
Josefa Ibañez, RN
Rubén Román, MSc
Francisco Ferrer, MD, PhD
Alfonso Vega, MD, PhD
Maria Soledad Laso, MD
Xavier Castells, MD, PhD

Purpose: To evaluate the effect of the introduction of digital mammography on the recall rate, detection rate, false-positive rate, and rates of invasive procedures in a cohort of women from four population-based breast cancer screening programs in Spain.

Materials and Methods: The study was approved by the ethics committee; informed consent was not required. A total of 242 838 mammograms (171 191 screen film [screen-film mammography group] and 71 647 digital [digital mammography group]) obtained in 103 613 women aged 45–69 years were included. False-positive results for any additional procedure and for invasive procedures, the breast cancer rate, and the positive predictive value in each group were compared by using Pearson χ^2 test. The effect of the mammographic technology used (screen-film or digital) on the false-positive results and cancer detection risk was evaluated with multivariate logistic regression models, adjusted according to women's and the screening program's characteristics and time trends.

Results: The false-positive rate was higher for screen-film than for digital mammography (7.6% and 5.7%, respectively; $P < .001$). False-positive results after an invasive procedure were significantly higher for screen-film than for digital mammography (1.9% and 0.7%, respectively; $P < .001$). No significant differences were observed in the overall cancer detection rate between the two groups (0.45% and 0.43% in the screen-film and digital mammography groups, respectively; $P = .59$). The adjusted risk of a false-positive result was higher for screen-film than for digital mammography (odds ratio = 1.32). The adjusted risk was also lower for the digital mammography group when time trends were taken into account.

Conclusion: The lower false-positive risk with use of digital mammography should be taken into account when balancing the risks and benefits of breast cancer screening.

[1] From the Health Services Evaluation and Clinical Epidemiology Department, IMIM-Hospital del Mar, CIBER Epidemiología y Salud Pública (CIBERESP), Passeig Marítim 25-29, 08008 Barcelona, Spain. Received April 30, 2010; revision requested June 7; revision received July 20; accepted July 27; final version accepted August 25. Supported by grants from the Health Ministry of Spain (Fondo de Investigación Sanitaria PI07/90293). **Address correspondence to** M.S. (e-mail: *Msalaserra@hospitaldelmar.cat*).

Digital technology is increasingly used for both diagnostic and screening mammography (1). Diagnostic performance of digital mammography (DM) is generally accepted to be at least equal to that of conventional screen-film mammography (SFM). However, studies comparing DM and SFM in screening mammography have shown divergent results, mainly in the recall rates, partly due to differences in study designs and several other factors. A recent review by Skaane (2) concluded that DM has an overall higher cancer detection rate than does SFM in screening mammography, achieved at the cost of a higher recall rate. However, a meta-analysis by Vinnicombe et al (3), including six of the seven studies analyzed in the review by Skaane, could not calculate pooled estimates for recall rates because they varied greatly among studies. The latest update of breast cancer screening by the U.S. Preventive Services Task Force (4) for the recent recommendation statement (5) included information on the benefits and harms of breast cancer screening based predominantly on studies of SFM but not on studies of DM.

Advances in Knowledge

- The recall rate, false-positive results overall, and false-positive findings resulting in an invasive procedure were lower with digital mammography (DM) with soft-copy reading (6.2%, 5,7%, and 0.74%, respectively) than with screen-film mammography (8.1%, 7.6%, and 1.9%, respectively) in four population-based screening programs in Spain ($P < .001$).
- The lower false-positive risk remained after adjustment was made for women's screening mammogram, age at screening, radiologic unit screening round, screening program characteristics, and time trends.
- No differences were observed in the cancer detection rate between screen-film and DM in women younger than 50 years of age or in the first or successive screening rounds.

Information on the effects of DM within operational population-based breast cancer screenings programs is still scarce.

Some authors have started to report results of the impact of DM in a population-based screening practice (6–9), but the results on recall rate are controversial. One of the main limitations of these studies was the relatively low number of screening tests performed with DM and the fact that most were performed within a single program and no adjustments were made to control for other risk factors affecting the recall rate.

Two of the main disadvantages of breast cancer screening are the recall rate and false-positive results, which, although intrinsic parts of the program, may lead to additional (sometimes invasive) tests, thus increasing costs and provoking anxiety in women before malignancy is ruled out (10). Therefore, determining whether DM increases or reduces recall rates and false-positive results, with similar diagnostic performance to that of SFM, is of great importance, especially at a time when the risks and benefits of screening mammography are being debated (11,12).

The aim of this study was to evaluate the effect of the introduction of DM on the recall rate, detection rate, false-positive rate, and rates of invasive procedures in a cohort of women from four population-based breast cancer screening programs in Spain.

Materials and Methods

Setting

Five radiology units from four different population-based breast cancer screening programs in Spain were enrolled, covering a population of 1 300 000 inhabitants.

Implications for Patient Care

- DM with soft-copy reading could prevent a number of women from recall and false-positive results without affecting the cancer detection rate.
- The lower false-positive risk with DM should be taken into account when balancing the risks and benefits of breast cancer screening.

The study was approved by the ethics committee, and informed consent was not required. The selection criterion for inclusion in the screening programs was having completed at least one screening round performed with DM by December 2007. A total of 19032 screening mammograms were included in a previous study (9). All programs are based on the European guidelines for quality assurance in screening mammography (13), and their results meet the required standards. Women in the target population receive information about screening and are invited to undergo mammography with a 2-year interval between screening rounds.

All radiology units began screening activities between 1996 and 1998 by using screen-film radiographic technology and switched to full-field DM between September 2004 (one program) and January 2005 (three programs). Two programs used a 2000D unit (Senographe; GE Medical Systems, Milwaukee, Wis) and two used a DM 1000 unit (Agfa; Lorad, Danbury, Conn).

By December 2007, one radiology unit had finished the fourth screening round, two units were at the fifth screening round, and two units were at the sixth screening round. The age at screening was 50–69 years in three programs and 45–69 years in one program. All

Published online
10.1148/radiol.10100874

Radiology 2011; 258:388–395

Abbreviations:
CI = confidence interval
DM = digital mammography
OR = odds ratio
PPV = positive predictive value
SFM = screen-film mammography

Author contributions:
Guarantors of integrity of entire study, M. Sala, D.S., X.C.; study concepts/study design or data acquisition or data analysis/interpretation, all authors; manuscript drafting or manuscript revision for important intellectual content, all authors; approval of final version of submitted manuscript, all authors; literature research, M. Sala, D.S., F.B., J.F., J.I., F.F., A.V., M.S.L., X.C.; clinical studies, M. Sala, M. Sánchez, J.F., F.F., A.V., M.S.L.; statistical analysis, M. Sala, F.B., M. Sánchez, R.R., X.C.; and manuscript editing, M. Sala, D.S., M. Sánchez, F.F., R.R., X.C.

Potential conflicts of interest are listed at the end of this article.

programs obtained two views (mediolateral oblique and craniocaudal) from 2007 onward. Before this date, a single view was obtained for subsequent screenings in one program. Reading methods were single reading in one program, double reading with consensus in two programs, and double reading with arbitration in one program. The different reading methods were used equally within each program before and after the introduction of the digital method. None of the programs used a computer-aided detection system.

Twenty-six radiologists participated in the study. Experience in mammogram reading varied, but most radiologists were involved in the programs from their beginning and all had read a minimum number of mammograms (between 3000 and 5000) before entering the screening program, as recommended by the European guidelines (13). The final database covered information on the women's age, mammographic technology (analog or digital), and further assessments after a positive screening mammogram. A definitive diagnosis of breast cancer was always confirmed histopathologically. No information about previous mammograms was provided at reading at the first screening round. Previous mammograms obtained with SFM were not digitized. All information was collected at each woman's attendance, and no major changes took place in the reading protocol when DM was introduced.

Study Population

This study included women participating in at least one screening in any of the four screening programs. Mammograms obtained with DM during the first 3 months after the technology was changed, which could be considered a learning period, were excluded and only mammograms with soft-copy reading were included in the analysis. A total of 242838 screening mammograms from 103613 women were included in the analysis, of which 171191 were screen film (SFM group) and 71647 were digital (DM group).

Screening Results

Two possible outcomes of a screening test were considered: normal findings (for which follow-up at 2 years is recommended) or positive results (abnormal findings requiring a recall for immediate further assessment to exclude malignancy). As proposed in the European guidelines (13), the program did not include an early recall as a direct result of the findings from screening mammogram, that is, recommendation for a woman to undergo a short-term rescreening at an interval shorter than the program's routine screening round length (2 years) without any other additional investigation. A positive result was considered a true-positive result if, after further assessments, breast cancer was found (ductal carcinoma in situ or invasive cancer). Otherwise, the result was considered to be false-positive. Further assessments could include both noninvasive (additional mammography, ultrasonography, magnetic resonance imaging) and invasive procedures (fine-needle aspiration cytologic analysis, core-needle biopsy, and open biopsy). False-positive results were classified into two types: false-positive overall (noninvasive and/or invasive, further assessment was performed) and false-positive resulting in an invasive procedure (at least one invasive, further assessment was performed). Early recalls were not considered as false-positive results if they did not involve further procedures that resulted in a noncancer. Only repeat screenings achieving sufficient technical quality were included as the screening examination.

Statistical Analysis

The recall rate was defined as the percentage of screened women requiring at least one further assessment after a positive mammogram. The detection rate was defined as the percentage of screened women with a true-positive result, that is, a final diagnosis of breast cancer (invasive cancer and ductal carcinoma in situ). The positive predictive value (PPV) of screening was defined as the fraction of recall examinations leading to a diagnosis of breast cancer.

The recall rate, overall false-positive results, false-positive findings resulting in an invasive procedure, the breast cancer rate, and the PPV were computed for each group by using simple proportions and were compared by using the Pearson χ^2 test and the Fisher exact test. Multivariate logistic regression models were constructed to evaluate the effect of the mammographic technology used (SFM and DM) on the false-positive results and the odds ratio (OR) of cancer detection, adjusted according to women's screening round (women who were undergoing their first to sixth screening round; the second to sixth screening rounds for each woman were included in the group of successive screening rounds), age at screening, the screening round of the radiology unit, and the screening program. The number of views and the reading method were not included because they are variable and highly associated with specific programs.

False-positive models included the screening program in which mammography was performed as a random effect, since we were concerned about the clustered structure of the data. To account for repeated measures in the same participant, we also included each woman's participations as a random effect (compound symmetry structure) (14). However, no random effects were included in the cancer detection models since no effect for the screening program was found (the covariance matrix for program effects was null), and repeated measures in the same woman were considered independent observations (cancer detection is always conditional on not having had previous breast cancer).

To exclude potential confounding factors due to time trends, we divided the screening history in each radiology unit into six consecutive time intervals. The SFM period was divided into four intervals with an equal number of mammograms (quartiles). Similarly, each DM period was divided into two equal intervals by using the median date of mammography. In each radiology unit, the first digital period was preceded by four analog periods. Thus, we obtained a combined time/technique variable, which we called "SFM/FFDM periods." The logistic regression models were replicated by using the variable of the combined "SFM/FFDM periods" instead of the mammographic technique and radiology unit screening round independently to control

for possible errors in the estimations due to a high correlation between them.

All calculations were performed by using software (SAS system for Windows, version 9.1.3; SAS Institute, Cary, NC). The logistic procedure was used for the cancer detection model and the Glimmix software (version 9.1; SAS Institute) was used for the false-positive model. An α level was set at .05, and all tests were two tailed.

Results

A total of 242 838 screening mammograms obtained in 103 613 women were included in the analysis. The overall recall rate was higher in the SFM group than in the DM group (13 860 [8.1%] of 171 191 and 4420 [6.2%] of 71 647, respectively; $P < .001$; Table 1). The recall rate at first screening round was higher in the SFM group than in the DM group (12.1% and 11.7%, respectively; $P = .091$), as well as at successive screening rounds (5.0% in SFM and 4.6% in DM group; $P < .001$). Early recalls were also higher in the SFM group than in the DM group (0.78% and 0.25%, respectively; $P < .001$). False-positive results, in agreement with recall rates, were higher in the SFM group than in the DM group (7.6% vs 5.7%, $P < .001$), but differences were not statistically significant at the first screening round (11.6% vs 11.1%; $P = .078$). False-positive findings resulting in an invasive procedure were significantly higher in the SFM group than in the DM group for both the first (3.0% vs 1.7%, respectively; $P < .001$) and successive screening rounds (1% vs 0.45%, respectively; $P < .001$). In total, 1080 cancers were detected, 770 in the SFM group and 310 in the DM group, representing a cancer detection rate of 0.45% in the SFM group and of 0.43% in the DM group ($P = .592$). No statistically significant differences were observed in cancer detection rates in the first screening round between the SFM and DM groups (0.52% vs 0.53%, respectively; $P = .862$) or at successive screening rounds (0.40% and 0.40%, respectively; $P = .834$). The percentage of ductal carcinoma in situ tumors was higher in the DM group for both the first (13.2% vs 17.4, $P = .005$) and successive (13.5% vs 18.8%, $P < .001$) screening rounds. The PPV was 5.6% (770 cancers in 13 860 recalled women) for SFM and 7% (310 of 4420) for DM.

The cancer detection risk was 4.8% higher with SFM than with DM (OR = 1.05; 95% confidence interval [CI]: 0.87, 1.29; Table 2). The first screening round had an increased risk of cancer detection (OR = 1.57; 95% CI: 1.23, 2.02) compared with the fourth or subsequent rounds. The cancer detection risk increased with age. The estimated OR in the group aged 45–49 years was 0.44 (95% CI: 0.32, 0.59) compared with the group aged 65–69 years. No significant differences were found between the groups aged 60–64 years and 65–69 years.

A significantly increased risk of a false-positive result overall was observed for SFM compared with DM (OR = 1.32; 95% CI: 1.25, 1.40; Table 3). The first, second, and third screening mammograms also had an increased risk of a false-positive result related to the fourth and subsequent screenings (OR = 2.95; 95% CI: 2.74, 3.17; first screening). Similarly, younger age groups were at higher risk of a false-positive result than the group aged 65–69 years (OR = 1.59; 95% CI: 1.48, 1.72; group aged 45–49 years). The risk of a false-positive finding resulting in an invasive procedure was higher for SFM than for DM (OR = 1.64; 95% CI: 1.43, 1.87) in the first screening round than in successive screening rounds (OR = 3.49; 95% CI: 2.95, 4.1) and in younger age groups (OR = 1.42; 95% CI: 1.20, 1.70; group aged 45–49 years).

The second and third SFM periods had a significantly lower risk of cancer detection compared with the first SFM period (second SFM period: OR = 0.73, 95% CI: 0.59, 0.91; third SFM period: OR = 0.78, 95% CI: 0.63, 0.97; Fig 1). The fourth SFM period and the first and second DM periods had lower cancer detection ORs compared with the first SFM period, although this difference was not statistically significant, and the OR for the second DM period was slightly higher than that for the previous period (OR = 0.87 and 0.82, respectively).

The false-positive risk for DM periods was significantly lower than that for the first SFM period (first DM period: OR = 0.81, 95% CI: 0.76, 0.86; second DM period: OR = 0.79, 95% CI: 0.74, 0.84; Fig 2). The risk of a false-positive finding resulting in an invasive procedure (Fig 3) was also lower in the DM periods than in the first SFM period (first DM period: OR = 0.61, 95% CI: 0.53, 0.71; second DM period: OR = 0.50, 95% CI: 0.42, 0.59).

Discussion

This study retrospectively analyzed a cohort of women from four breast cancer screening programs in which FSM and/or DM were used between 1996 and 2007. The results of these analyses show that while cancer detection did not differ in women screened with SFM or DM, the recall rate and false-positive risk were lower with DM than with SFM after adjustment was made for women's screening mammogram, radiology unit screening round, age at screening, and time trends.

These results seem to contradict the final conclusion of a review by Skaane (2). One of the limitations of that review, as well as of the meta-analysis by Vinnicombe et al (3), is that they included few studies performed in real population-based breast cancer screening conditions and the studies included differed fairly widely in their characteristics. Because DM was introduced in the period 2000–2005, the number of DMs in all of the studies was relatively small, being fewer than 20 000 in all studies (3,6,7) except one, which included 26 000 DMs (8). One of the studies (3) used hard-copy reading, while others used soft-copy reading (6,7). Moreover, only two studies included initial and successive screening mammograms, and only one (3) assessed the interaction between the screening round and the technique. A learning period after the introduction of the digital technique was only excluded in one study (8). Only one of the studies performed a multivariate analysis (3), allowing simultaneous control of distinct variables that could have had an effect on the detection rate, recall

Table 1

Screening Performance Indicators in SFM and DM Groups

Parameter	SFM*	DM*	*P* Value†	Total*
No. of screened women‡	84871	61795	…	103613
No of screening tests	171191 (70.5)	71647 (29.5)	…	242838 (100)
Recall	13860 (8.1)	4420 (6.2)	<.001	18280 (7.5)
First screening	8995 (12.1)	1886 (11.6)	.091	10881 (12.0)
Successive	4865 (5.0)	2534 (4.6)	<.001	7399 (4.9)
Early recall	1337 (0.78)	179 (0.25)	<.001	1516 (0.62)
First screening	1016 (1.4)	72 (0.44)	<.001	1088 (1.2)
Successive	321 (0.33)	107 (0.19)	<.001	428 (0.28)
False-positive for any procedure	13090 (7.6)	4110 (5.7)	<.001	17200 (7.1)
First screening	8609 (11.6)	1800 (11.1)	.078	10409 (11.5)
Successive	4481 (4.6)	2310 (4.2)	<.001	6791 (4.5)
False-positive for invasive procedures	3215 (1.9)	529 (0.74)	<.001	3744 (1.5)
First screening	2248 (3.0)	282 (1.7)	<.001	2530 (2.8)
Successive	967 (1.0)	247 (0.45)	<.001	1214 (0.80)
Cancer Detection	770 (0.45)	310 (0.43)	.592	1080 (0.44)
Age <50 y	48 (0.31)	22 (0.32)	.921	70 (0.31)
Age ≥50 y	722 (0.46)	288 (0.44)	.552	1010 (0.46)
First screening	386 (0.52)	86 (0.53)	.862	472 (0.52)
Ductal carcinoma in situ	51 (13.2)	15 (17.4)	…	66 (14.0)
Invasive cancer§	332 (86.0)	66 (76.7)	…	398 (84.3)
T1 (<10 mm)	190 (65.1)	44 (66.7)	…	234 (65.3)
T2 (10–20mm)	67 (22.9)	17 (25.7)	…	84 (23.4)
T3 (>20 mm)	5 (1.7)	1 (1.5)	.527	6 (1.7)
T4	9 (3.1)	0 (0)	…	9 (2.5)
Unknown	21 (7.2)	4 (6.0)	…	27 (7.5)
Total	292	66	…	358
Unknown	3 (0.78)	5 (5.8)	…	8 (1.7)
Successive screening	384 (0.40)	224 (0.40)	.834	608 (0.40)
Ductal carcinoma in situ	52 (13.5)	42 (18.8)	…	94 (15.5)
Invasive cancer‖	332 (86.5)	174 (77.7)	…	506 (83.2)
T1 (<10 mm)	211 (68.3)	110 (71.0)	…	321 (69.2)
T2 (10–20mm)	56 (18.1)	20 (12.9)	…	76 (16.4)
T3 (>20 mm)	8 (2.6)	2 (1.3)	.179	10 (2.2)
T4	5 (1.6)	0 (0)	…	5 (1.1)
Unknown	29 (9.4)	23 (14.8)	…	52 (11.2)
Total	309	155	…	464
Unknown	0 (0.00)	8 (3.6)	…	8 (1.3)
PPV	770/13860 (5.6)	310/4420 (7.0)	<.001	1080/18280 (5.9)
First screening	386/8995 (4.3)	86/1886 (4.6)	.663	472/10881 (4.3)
Successive screening	384/4865 (7.9)	224/2534 (8.8)	.211	608/7399 (8.2)

* Data are number of patients; data in parentheses are percentages.

† Fisher exact test.

‡ The total number of screened women is not the sum of women screened with SFM and DM because on the basis of their screening history, both techniques were used in 43053 women.

§ No information available about size and stage for 40 invasive tumors.

‖ No information available about size and stage for 42 invasive tumors.

rate, and false-positive results, such as the screening round, patient age, and reading protocol characteristics.

In a previous study performed in one of the programs included in the present study (9), two screening rounds of a population-based breast cancer screening program (one round before and one after the implementation of digital technology) were compared and no significant differences were observed in the cancer detection rate between the two screening rounds, while the recall rates and rates of invasive tests were lower in the DM group. Results in the same direction were observed in our study, with no statistically significant differences being observed in the cancer detection

rate between DM and SFM, either in the first or in successive screening rounds; however, the detection rate was higher with DM than with SFM in the first screening round, which usually includes the youngest women. Pisano et al (15,16) found that diagnostic accuracy was better with DM in women younger than 50 years, with dense breast tissue, and pre-or perimenopausal status. This finding was not observed in the present study, where no statistically significant differences were observed in the cancer detection rate according to women's age. Because no information about the breast density was available in the screening programs included in this study, we could not assess the effect of this variable. Some studies have found higher rates of ductal carcinoma in situ with DM, which could be partly due to the improved detection of microcalcifications with DM, but we did not analyze this possibility. In addition, a decrease in invasive breast cancer, mainly in 2001 in women aged 45–64 years, was described in Spain (17), which was explained in part by the screening saturation effect and the decline in the use of hormone replacement therapy.

The recall rate, false-positive results overall, and false-positive findings resulting in an invasive procedure were found to be lower with DM than with SFM and remained lower after adjustment was made for women's age, women's screening mammogram, radiology unit screening round (programs started in different years), and the program to control for the variability introduced by differences in the programs' characteristics such as the number of views and the reading method. A reduction in the recall rate in the DM group was reported in one paired study (18) and by two studies performed in a daily-practice environment (8,9), while no differences in the recall rate were observed between the two modalities in two further studies performed in a daily-practice environment (3,6) and in one paired study (15). Although we could not control for radiologists' experience or some other conditions such as mammography viewing, the analysis controlling for period and technique suggested that the reduction in false-positive results was not due to time trends associated with other factors such as viewing conditions, greater experience, or the reduction in the number of hormone replacement therapy users after the Women's Health study (19) results were published. All these differences and uncontrolled

Table 2

Logistic Regression Model for Cancer Detection*

Variable	Cancer		Crude OR†	Adjusted OR*
	No	Yes		
Technology				
SFM	170421	770	1.040 (0.911, 1.187)	1.048 (0.847, 1.298)
Full-field DM	71337	310	1.000	1.000
Women's screening mammogram				
1	89908	472	1.196 (1.003, 1.426)	1.579 (1.233, 2.022)
2	65942	242	0.836 (0.687, 1.018)	1.194 (0.937, 1.522)
3	47184	196	0.946 (0.770, 1.163)	1.215 (0.959, 1.539)
≥4	38724	170	1.000	1.000
Age at screening (y)				
45–49	22220	70	0.561 (0.426, 0.739)	0.440 (0.324, 0.597)
50–54	64882	247	0.678 (0.561, 0.820)	0.554 (0.451, 0.680)
55–59	62729	244	0.693 (0.573, 0.838)	0.614 (0.506, 0.747)
60–64	58250	330	1.009 (0.844, 1.208)	0.897 (0.746, 1.079)
65–69	33677	189	1.000	1.000
Radiology unit screening round				
1st	43207	267	1.373 (1.177, 1.601)	1.053 (0.806, 1.377)
2nd	53895	204	0.841 (0.711, 0.994)	0.753 (0.583, 0.972)
3rd	51347	189	0.818 (0.689, 0.971)	0.738 (0.585, 0.931)
≥4th	93309	420	1.000	1.000

* Model adjusted according to variables in the Table and program. Each mammogram has been considered an independent observation.

† Data in parentheses are 95% CIs.

Figure 1

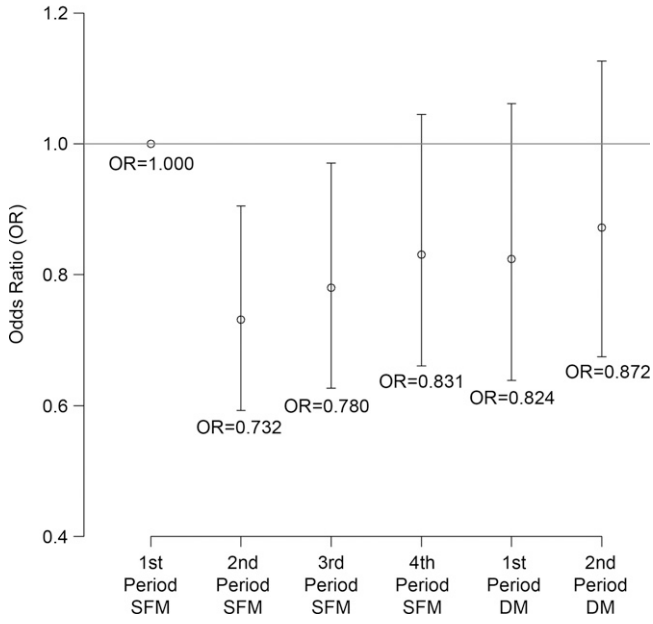

Figure 1: Adjusted OR for cancer detection in SFM and DM periods.

Table 3

Logistic Regression Model for False-Positive Results*

	Invasive and Noninvasive Procedures				Invasive Procedures			
	False-positive				False-positive			
Variable	No	Yes	Crude OR†	Adjusted OR†	No	Yes	Crude OR†	Adjusted OR*†
Technology								
SFM	158101	13090	1.361 (1.312, 1.411)	1.325 (1.247, 1.407)	167976	3215	2.573 (2.346, 2.822)	1.639 (1.436, 1.870)
Full-field DM	67537	4110	1.000	1.000	71118	529	1.000	1.000
Women's screening mammogram								
1	79971	10409	3.782 (3.565, 4.012)	2.951 (2.741, 3.177)	87850	2530	4.779 (4.176, 5.468)	3.487 (2.958, 4.111)
2	62780	3404	1.576 (1.476, 1.682)	1.304 (1.209, 1.405)	65574	610	1.544 (1.327, 1.796)	1.263 (1.065, 1.497)
3	45287	2093	1.343 (1.251, 1.441)	1.219 (1.130, 1.317)	47009	371	1.310 (1.111, 1.543)	1.163 (0.976, 1.386)
4	37600	1294	1.000	1.000	38661	233	1.000	1.000
Age at screening (y)								
45–49	19424	2866	3.323 (3.111, 3.548)	1.595 (1.479, 1.721)	22009	281	1.252 (1.068, 1.467)	1.425 (1.190, 1.706)
50–54	59283	5846	2.221 (2.093, 2.356)	1.427 (1.338, 1.521)	63691	1438	2.213 (1.966, 2.492)	1.306 (1.147, 1.487)
55–59	59153	3820	1.454 (1.367, 1.547)	1.137 (1.066, 1.213)	62092	881	1.391 (1.227, 1.577)	1.023 (0.898, 1.166)
60–64	55352	3228	1.313 (1.232, 1.399)	1.037 (0.972, 1.106)	57778	802	1.361 (1.198, 1.545)	0.948 (0.833, 1.080)
65–69	32426	1440	1.000	1.000	33524	342	1.000	1.000
Radiology unit screening round								
1st	38433	5041	2.219 (2.131, 2.311)	0.904 (0.844, 0.969)	42242	1232	2.713 (2.493, 2.951)	1.055 (0.932, 1.195)
2nd	50226	3873	1.305 (1.250, 1.362)	0.886 (0.829, 0.947)	53226	873	1.526 (1.392, 1.672)	0.949 (0.840, 1.073)
3rd	48481	3055	1.066 (1.018, 1.116)	0.888 (0.834, 0.945)	50894	642	1.173 (1.062, 1.296)	0.963 (0.853, 1.086)
≥4th	88498	5231	1.000	1.000	92732	997	1.000	1.000

* Model adjusted according to variables in the Table and program. "Program" was entered as a G-side random effect and "women" was entered as an R-side random effect.

† Data in parentheses are 95% CIs.

Figure 2

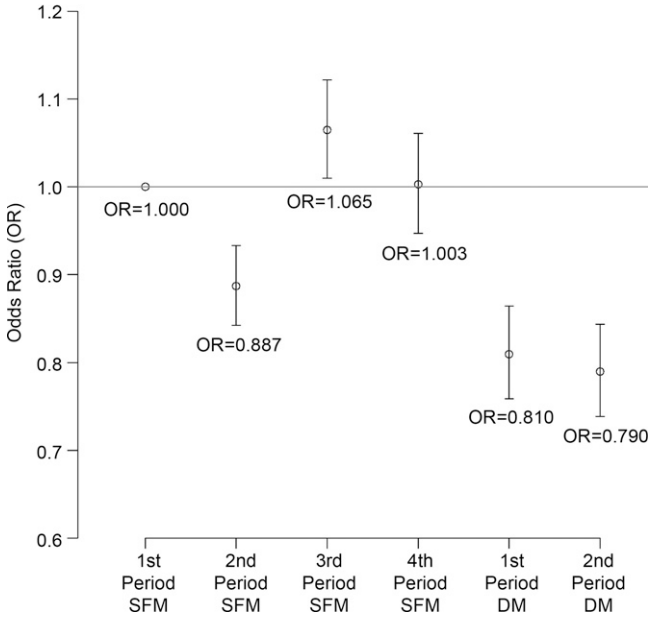

Figure 2: Adjusted OR for false-positive results overall in SFM and DM periods.

Figure 3

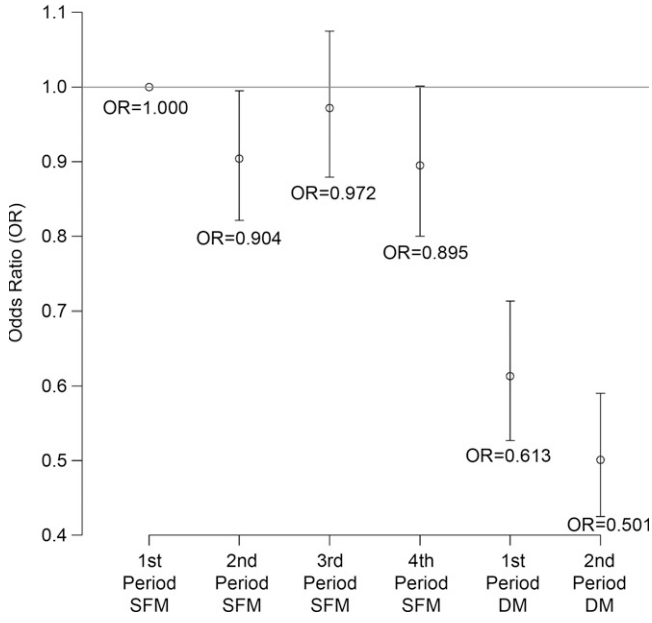

Figure 3: Adjusted OR for false-positive findings resulting in an invasive procedure in SFM and DM periods.

effects partly explained the differences observed in the results on the recall rate, as discussed elsewhere (1,20), and suggest that factors other than the mammographic technique may play a major role in the risk of recall and false-positive results. However, the real impact of DM needs to be understood for complete evaluation of the adverse effects of breast cancer screening.

False-positive rates after invasive procedures were lower with DM than with SFM. We found no studies comparing the results of invasive tests between the two modalities. Nevertheless, the adverse effect of a false-positive result after an invasive procedure is higher in terms of the physical impact to women and involves a higher cost than imaging procedures and a delay in informing women of the results.

One of the limitations of this study was the relatively short period for which there is experience with DM and that many factors that were not controlled may have influenced the quality and process indicators. For instance, PPV is affected by prevalence, which was not controlled for in the present study. However, we did control for some important factors such as time trends, women's screening mammogram, age at screening, and radiology unit screening round and included information from different breast cancer screening programs.

Although information about false-negative results from women in the DM group is lacking, which hampered complete evaluation of the impact of the introduction of this technique, these results suggest that the introduction of DM, apart from its technical advantages, does not increase the adverse effects related to recall rate and false-positive results. On the contrary, reduction of the false-positive rate could prevent a very large number of women from experiencing the consequences of this adverse effect.

Acknowledgments: We are grateful to Francesc Macià, MD, Sonia Sánchez, MD, Paula Merino, MD, and Cristina Hernández for their contribution in the data collection of the screening programs.

Disclosures of Potential Conflicts of Interest: M.Sala. Financial activities related to the present article: Received a grant from Health Ministry of Spain (PI07/90293). Financial activities not related to the present article: none to disclose. Other relationships: none to disclose. **DS.** Financial activities related to the present article: Received a grant from Health Ministry of Spain (PI07/90293). Financial activities not related to the present article: none to disclose. Other relationships: none to disclose. **F.B.** Financial activities related to the present article: Received a grant from Health Ministry of Spain (PI07/90293). Financial activities not related to the present article: none to disclose. Other relationships: none to disclose. **M. Sánchez.** Financial activities related to the present article: Received a grant from Health Ministry of Spain (PI07/90293). Financial activities not related to the present article: none to disclose. Other relationships: none to disclose. **J.F.** Financial activities related to the present article: Received a grant from Health Ministry of Spain (PI07/90293). Financial activities not related to the present article: none to disclose. Other relationships: none to disclose. **M.J.I.** Financial activities related to the present article: Received a grant from Health Ministry of Spain (PI07/90293). Financial activities not related to the present article: none to disclose. Other relationships: none to disclose. **R.R.** Financial activities related to the present article: Received a grant from Health Ministry of Spain (PI07/90293). Financial activities not related to the present article: none to disclose. Other relationships: none to disclose. **F.F.** Financial activities related to the present article: Received a grant from Health Ministry of Spain (PI07/90293). Financial activities not related to the present article: none to disclose. Other relationships: none to disclose. **A.V.** Financial activities related to the present article: Received a grant from Health Ministry of Spain (PI07/90293). Financial activities not related to the present article: none to disclose. Other relationships: none to disclose. **M.S.L.** Financial activities related to the present article: Received a grant from Health Ministry of Spain (PI07/90293). Financial activities not related to the present article: none to disclose. Other relationships: none to disclose. **X.C.** Financial activities related to the present article: Received a grant from the Health Ministry of Spain (PI07/90293). Financial activities not related to the present article: none to disclose. Other relationships: none to disclose.

References

1. Bick U, Diekmann F. Digital mammography: what do we and what don't we know? Eur Radiol 2007;17(8):1931–1942.
2. Skaane P. Studies comparing screen-film mammography and full-field digital mammography in breast cancer screening: updated review. Acta Radiol 2009;50(1):3–14.
3. Vinnicombe S, Pinto Pereira SM, McCormack VA, Shiel S, Perry N, Dos Santos Silva IM. Full-field digital versus screen-film mammography: comparison within the UK breast screening program and systematic review of published data. Radiology 2009;251(2):347–358.
4. Nelson HD, Tyne K, Naik A, et al. Screening for breast cancer: an update for the U.S. Preventive Services Task Force. Ann Intern Med 2009;151(10):727–737, W237–W42.
5. US Preventive Services Task Force. Screening for breast cancer: U.S. Preventive Services Task Force recommendation statement. Ann Intern Med 2009;151(10):716–726, W236. [Published correction appears in Ann Intern Med 2010;152(3):199–200.]
6. Vigeland E, Klaasen H, Klingen 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(1):183–191.
7. Del Turco MR, Mantellini P, Ciatto S, et al. Full-field digital versus screen-film mammography: comparative accuracy in concurrent screening cohorts. AJR Am J Roentgenol 2007;189(4):860–866.
8. Heddson 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(3):419–425.
9. Sala M, Comas M, Macià F, Martinez J, Casamitjana M, Castells X. Implementation of digital mammography in a population-based breast cancer screening program: effect of screening round on recall rate and cancer detection. Radiology 2009;252(1):31–39.
10. Brewer NT, Salz T, Lillie SE. Systematic review: the long-term effects of false-positive mammograms. Ann Intern Med 2007;146(7):502–510.
11. Murphy AM. Mammography screening for breast cancer: a view from 2 worlds. JAMA 2010;303(2):166–167.
12. Peres J. Mammography screening: after the storm, calls for more personalized approaches. J Natl Cancer Inst 2010;102(1):9–11.
13. Perry N, Broeders M, deWolf C, Tönberg S, Holland R, von Karsa L, eds. European guidelines for quality assurance in breast cancer screening and diagnosis. Luxembourg: European Commission, 2006.
14. Littell RC, Milliken GA, Stroup WW, Wolfinger RD, Schabenberger O. SAS for mixed models. 2nd ed. Cary, NC: SAS Institute, 2006.
15. Pisano ED, Gatsonis C, Hendrick E, et al. Diagnostic performance of digital versus film mammography for breast-cancer screening. N Engl J Med 2005;353(17):1773–1783.
16. Pisano ED, Hendrick RE, Yaffe MJ, et al. Diagnostic accuracy of digital versus film mammography: exploratory analysis of selected population subgroups in DMIST. Radiology 2008;246(2):376–383.
17. Pollán M, Pastor-Barriuso R, Ardanaz E, et al. Recent changes in breast cancer incidence in Spain, 1980-2004. J Natl Cancer Inst 2009;101(22):1584–1591.
18. Lewin JM, Hendrick RE, D'Orsi CJ, et al. Comparison of full-field digital mammography with screen-film mammography for cancer detection: results of 4945 paired examinations. Radiology 2001;218(3):873–880.
19. Barbaglia G, Macià F, Comas M, et al. Trends in hormone therapy use before and after publication of the Women's Health Initiative trial: 10 years of follow-up. Menopause 2009;16(5):1061–1064.
20. 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(3):708–717.