Radiology

Influence of Risk Category and Screening Round on the Performance of an MR Imaging and Mammography Screening Program in Carriers of the *BRCA* Mutation and Other Women at Increased Risk[1]

Suzan Vreemann, MS
Albert Gubern-Mérida, PhD
Margrethe S. Schlooz-Vries, MD
Peter Bult, MD, PhD
Carla H. van Gils, PhD
Nicoline Hoogerbrugge, MD, PhD
Nico Karssemeijer, PhD
Ritse M. Mann, MD, PhD

Purpose: To evaluate the real-life performance of a breast cancer screening program for women with different categories of increased breast cancer risk with multiple follow-up rounds in an academic hospital with a large screening population.

Materials and Methods: Screening examinations (magnetic resonance [MR] imaging and mammography) for women at increased breast cancer risk (January 1, 2003, to January 1, 2014) were evaluated. Risk category, age, recall for workup of screening-detected abnormalities, biopsy, and histopathologic diagnosis were recorded. Recall rate, biopsy rate, positive predictive value of recall, positive predictive value of biopsy, cancer detection rate, sensitivity, and specificity were calculated for first and follow-up rounds.

Results: There were 8818 MR and 6245 mammographic examinations performed in 2463 women. Documented were 170 cancers; of these, there were 129 screening-detected cancers, 16 interval cancers, and 25 cancers discovered at prophylactic mastectomy. Overall sensitivity was 75.9% including the cancers discovered at prophylactic mastectomy (95% confidence interval: 69.5%, 82.4%) and 90.0% excluding those cancers (95% confidence interval: 83.3%, 93.7%). Sensitivity was lowest for carriers of the *BRCA1* mutation (66.1% and 81.3% when including and not including cancers in prophylactic mastectomy specimens, respectively). Specificity was higher at follow-up (96.5%; 95% confidence interval: 96.0%, 96.9%) than in first rounds (85.1%; 95% confidence interval: 83.4%, 86.5%) and was high for both MR imaging (97.1%; 95% confidence interval: 96.7%, 97.5%) and mammography (98.7%; 95% confidence interval: 98.3%, 99.0%). Positive predictive value of recall and positive predictive value of biopsy were lowest in women who had only a family history of breast cancer.

Conclusion: Screening performance was dependent on risk category. Sensitivity was lowest in carriers of the *BRCA1* mutation. The specificity of high-risk breast screening improved at follow-up rounds.

Online supplemental material is available for this article.

[1] From the Departments of Radiology and Nuclear Medicine (S.V., A.G.M., N.K., R.M.M.), Surgery (M.S.S.V.), Pathology (P.B.), and Human Genetics (N.H.), Radboud University Medical Centre, Geert Grooteplein 10, 6525 GA Nijmegen, the Netherlands; and Julius Center for Health Sciences and Primary Care, University Medical Center Utrecht, Utrecht, the Netherlands (C.H.v.G.). Received February 24, 2017; revision requested April 21; revision received June 28; accepted July 23; final version accepted August 14. **Address correspondence to** S.V. (e-mail: *Suzan.Vreemann@radboudumc.nl*).

Study supported by the Netherlands Organisation for Health Research and Development (90514524) and European Union's Seventh Framework Programme for research, technological development and demonstration (601040).

Breast magnetic resonance (MR) imaging is considered to be the most sensitive imaging modality for early breast cancer detection and is recommended as supplemental screening technique for women with a lifetime risk for the development of breast cancer of 20%–25% or higher (1,2). This includes women with a *BRCA* germline mutation, for whom the lifetime risk is as high as 56%–84%, and women who underwent radiation therapy to the chest wall at young age (1–5). Furthermore, women with a strong family history for breast cancer, women with a personal history of breast cancer, and women with high-risk lesions such as atypical ductal hyperplasia and lobular carcinoma in-situ are at increased risk, though for the latter risk categories the indication for additional screening with MR imaging is less clear (1,2,5). Nonetheless, supplemental screening may be indicated for all these women because the rate of interval cancers is relatively high (6–8).

Whereas the sensitivity of breast MR imaging has been reported (9) to be as high as 95% in women with known breast cancer, the sensitivity of breast MR imaging in screening is lower. Initially, the sensitivity was reported to be around 77% in the screening setting, but more recent studies document a higher sensitivity of around 90% (10–13). Studies (11–13) show a nearly doubled cancer detection of combined MR imaging and mammography screening compared with mammography screening alone. Consequently, high-risk screening programs with MR imaging and mammography have been implemented in clinical practices worldwide, although there are national, regional, and local differences regarding whom to screen with supplemental MR imaging (1,2,14). This is likely also because it is so far unknown whether the diagnostic value of breast MR imaging screening is different for groups with a different underlying risk and whether the added value of MR imaging persists in follow-up evaluations for all groups.

A major criticism regarding breast MR imaging screening, beyond availability and cost, is that the specificity appears lower than generally accepted for mammography screening. Reported specificities range from 81% to 95% and compare poorly to a specificity of 99% for mammography screening (4,10,12,13,15–17). However, to our knowledge, hardly any data exist on the effect of routine use of breast MR imaging and the availability of prior examinations on the diagnostic accuracy of breast MR imaging. Only a few recent studies of breast MR imaging screening programs document recall rates, false-positive findings, and interval cancers (11,13).

The purpose of this study is to evaluate the real-life performance of a breast cancer screening program for women with different categories of increased breast cancer risk with multiple follow-up rounds in an academic hospital with a large screening population.

Advances in Knowledge

- The risk category for screening is an important factor for the performance of a breast cancer screening program; positive predictive value of biopsy ranged between 0.14 and 0.39 and was lowest for women with only a family history of breast cancer.
- The sensitivity of a combined annual MR examination with mammography breast cancer screening program is lowest for carriers of the *BRCA1* mutation (81.3%, excluding cancers found in a specimen from prophylactic mastectomy).
- The specificity of the breast cancer screening program strongly improved in follow-up rounds and was much higher in clinical practice than reported in initial prospective studies (first round, 85.1%; follow-up, 96.5%).

Implications for Patient Care

- Knowledge of the performance of an annual breast cancer screening program with mammography and MR imaging allowed for clear patient information regarding potential benefits and harms of screening.
- The overall specificity of the program, especially in follow-up rounds, was high.

Materials and Methods

Screening Program

The breast cancer screening program for women at increased breast cancer risk (≥20% to 25% lifetime risk) in our hospital consists of annual breast MR imaging, starting from age 25 years in carriers of the *BRCA* mutation. From age 30 years or older, MR imaging is combined with a yearly mammographic examination in *BRCA* mutation carriers. In women at high familial risk, screening starts between 35 and 45 years or 5 years before the age at which the youngest relative developed breast cancer (18). In women with a personal history of breast cancer, screening starts 1 year after the diagnosis of breast cancer.

Case Selection

This retrospective study was approved by our local institutional review board and the requirement for informed consent was waived. The local database was searched to identify all screening MR examinations and mammographies performed from January 1, 2003, to

https://doi.org/10.1148/radiol.2017170458

Content code: BR

Radiology 2018; 286:443–451

Abbreviation:
BI-RADS = Breast Imaging Reporting and Data System

Author contributions:
Guarantors of integrity of entire study, S.V., N.H., R.M.M.; 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; agrees to ensure any questions related to the work are appropriately resolved, all authors; literature research, S.V., A.G.M., M.S.S.V., N.H., R.M.M.; clinical studies, S.V., M.S.S.V., N.H., R.M.M.; experimental studies, S.V., A.G.M., M.S.S.V.; statistical analysis, S.V., A.G.M., M.S.S.V., C.H.v.G.; and manuscript editing, all authors

Conflicts of interest are listed at the end of this article.

January 1, 2014. Women were included when MR examinations were performed on the basis of a screening indication, mammographic examinations were included if the indication was screening, and mammography was performed within 1 year after the screening MR examination. We recorded risk category (*BRCA1*, *BRCA2*, family history, personal history, and other women), age, screening tests performed, recall for workup of screening-detected abnormalities, biopsies, and histopathologic diagnosis when available.

During the study period, 9571 screening breast MR examinations and 6553 screening mammographic examinations were performed in 2773 women. After excluding women who were not at increased risk, we evaluated 8818 screening breast MR examinations and 6245 screening mammographic examinations that were performed in 2463 women at increased risk for breast cancer (mean age, 44 years ± 12 [standard deviation]; age range, 15–91 years). The population included 770 women who were *BRCA* mutation carriers (471 carriers of *BRCA1* and 299 carriers of *BRCA2*) and 26 untested first-degree relatives (16 carriers of *BRCA1* and 10 carriers of *BRCA2*). Furthermore, women were included on the basis of a family history of breast cancer (lifetime risk, >20%; n = 748), a personal history of breast cancer (n = 836), or other reasons (chest radiation and high-risk lesions, n = 83). The mean number of MR examinations performed per woman was three (range, one to 13 examinations), and the mean number of mammographic examinations was two (range, zero to 12 examinations). Characteristics of the study population are presented in Table 1.

Image Acquisition

MR imaging protocols varied over time. Examinations were performed on a 1.5- or 3.0-T imager (Magnetom Avanto, Magnetom Sonata, Magnetom Simphony, or Magnetom Trio; Siemens, Erlangen, Germany) by using a dedicated bilateral breast coil. Women were imaged in the prone position. A transverse or coronal three-dimensional T1-weighted gradient-recalled echo dynamic sequence was performed before contrast agent administration followed by four or five postcontrast-administration sequences. Pixel spacing, section thickness, matrix, echo time, repetition time, and flip angle differed among acquisitions as previously described by Dalmış et al (19). Diffusion-weighted imaging and T2-weighted acquisitions were added from those imaged in 2012. Various gadolinium chelates were used, administered at a dose of 0.1 mmol/kg or 0.2 mmol/kg by using a power injector (Medrad, Warrendale, Pa) and at a flow rate of 2.5 mL/sec, followed by a saline flush. Premenopausal women were scheduled between the 6th and 12th day of their menstrual cycle.

Mammograms were obtained in two directions (mediolateral oblique and craniocaudal) with a full-field digital mammography imager (GE Senograph 2000 or GE Senograph DS; GE, Fairfield, Conn). Additional views and spot-compression views were performed at request of the evaluating radiologist.

Image Interpretation

All examinations were evaluated by one of eight breast radiologists (R.M.M. and other radiologists) by using the Breast Imaging Reporting and Data System (BI-RADS) (20,21). Experience in breast MR imaging ranged from 0.5 to 23 years. For MR imaging, both morphologic and dynamic parameters were evaluated by using a dedicated breast MR imaging workstation (versions of DynaCAD, Invivo, Gainesville, Fla; and Philips, Best, the Netherlands). When available, mammograms, MR images, and previous examinations were evaluated together. Final assessment categories were given as specified in the American College of Radiology BI-RADS atlas (20). Biopsies were performed for lesions classified as BI-RADS category 4 and BI-RADS category 5, and a subset of lesions classified as BI-RADS category 3. The remaining BI-RADS category 3 lesions underwent short-term follow-up.

Ground Truth

Our database was linked to the nationwide population-based Netherlands Cancer Registry. Normal or benign screening examinations were confirmed by at least 1 year of clinical follow-up and regarded to be true-negative findings when no cancer was detected before the subsequent screening round. When no biopsy was indicated at short-term follow-up, at least 1 year of clinical follow-up was required to confirm benignity. All lesions that underwent biopsy and prophylactic mastectomies performed in-house were identified by a cross-computer search with our pathologic database. We subsequently recorded whether the biopsy was performed for screening findings or symptoms.

Data Analysis

To evaluate the performance, we defined performance metrics (per examination) in analogy to those used for evaluation of mammography screening programs (22).

Screening examinations negative for cancer were defined as screening examinations that were rated as BI-RADS category 1 or 2. True-negative screening examinations were defined as screening examinations negative for cancer in women in whom no cancer was detected in the subsequent year.

Results of pathologic analysis were grouped into malignant (ie, ductal carcinoma in situ, invasive, and metastatic cancer) and benign (all other findings) lesions. Screening-detected cancers were defined as cancers diagnosed after diagnostic workup initiated by screening findings. In the high-risk screening program we defined two groups of non–screening-detected cancers. The first group consisted of interval cancers that manifested in between screening rounds because of symptoms. The second group consisted of cancers that were detected in asymptomatic women who underwent prophylactic mastectomy. Prognostic differences between these groups were investigated by using χ^2 tests for categorical variables or one-way analysis of variance for continuous variables.

Consequently, two different groups of false-negative examinations (referred to here as FN1 and FN2) were defined.

Table 1

Characteristics of the Study Population

Parameter	Carriers of *BRCA1* (*n* = 471)	Untested First-degree Relatives *BRCA1* (*n* = 16)	Carriers of *BRCA2* (*n* = 299)	Untested First-degree Relatives *BRCA2* (*n* = 10)	Family History of Breast Cancer (*n* = 748)	Personal History of Breast Cancer (*n* = 836)	Other Patients at Increased Risk for Breast Cancer (*n* = 83)
Mean age at start of screening (y)	39 (23–75)	31 (22–69)	41 (23–73)	29 (25–50)	42 (16–73)	52 (22–91)	38 (15–71)
Median no. of screening MR imaging examinations performed	3 (1–13)	2 (1–4)	3 (1–11)	1 (1–10)	2 (1–12)	2 (1–12)	3 (1–10)
Median no. of screening mammography examinations performed	2 (0–11)	1 (0–5)	3 (0–11)	1 (0–10)	1 (0–11)	1 (0–12)	1 (0–7)
No. of available VDGs at start of screening*	398	11	265	7	526	474	52
Grade a	88 (22.1)	1 (9.1)	52 (19.6)	0 (0.0)	46 (8.7)	59 (12.4)	10 (19.2)
Grade b	104 (26.1)	1 (9.1)	63 (23.8)	2 (28.6)	83 (15.8)	130 (27.4)	10 (19.2)
Grade c	117 (29.4)	4 (36.4)	102 (38.5)	1 (14.3)	202 (38.4)	187 (39.5)	21 (40.4)
Grade d	89 (22.4)	5 (45.5)	48 (18.1)	4 (57.1)	195 (37.1)	98 (20.7)	11 (21.2)

Note.—Except where otherwise indicated, data in parentheses are range. VDG = volpara density grade.

* Volpara density grade is a volumetric and automatically assessed density score based on raw mammography data. The classes are created analogous to the BI-RADS lexicon. Data in parentheses are percentages.

FN1 consisted of screening examinations rated as BI-RADS category 1 or BI-RADS category 2 in which any cancer was detected in the following year or before the subsequent screening examination. FN2 consisted of screening examinations rated as BI-RADS category 1 or BI-RADS category 2 in which a true interval cancer was detected.

The recall rate was defined as the number of examinations that were classified as BI-RADS category 0, BI-RADS category 3, BI-RADS category 4, or BI-RADS category 5 per 1000 screening examinations performed. The biopsy rate was defined as the number of examinations in which a biopsy was performed per 1000 screening examinations performed. The cancer detection rate was defined as the number of examinations that led to breast cancer detection per 1000 screening examinations.

Screening examinations positive for cancer were defined as screening examinations that were rated as BI-RADS category 0, BI-RADS category 3, BI-RADS category 4, or BI-RADS category 5, and true-positive examinations were defined as screening examinations positive for cancer that led to cancer detection. False-positive examinations were defined as examinations that led to a recall in women in whom no breast cancer was detected. The positive predictive value of recall was defined as the fraction of recalls that led to cancer detection. The positive predictive value of biopsy was defined as the fraction of biopsies that led to cancer detection.

We calculated the sensitivity of the complete screening program and for both modalities separately. Three different sensitivities were recognized: the sensitivity for all breast cancers detected, the sensitivity for screening-detected cancers and interval cancers detected because of symptoms (excluding cancers detected in prophylactic mastectomies), and the modality-specific sensitivity for screening-detected cancers. For the latter category, lesions were considered visible with a modality when they were mentioned in the radiologic report of that modality (or report section when both modalities were reported simultaneously). Cancers only reported at mammography were considered to be false-negative findings at MR imaging and vice versa. Specificity was determined by using examinations of women that had at least 1 year of follow-up without recall and calculated for complete screening and independently for each modality.

We analyzed the overall screening results and separated results of the first screening round from follow-up rounds. The first screening round was defined as the first time a woman with increased breast cancer risk underwent MR imaging. Follow-up screening rounds were defined as screening examinations performed in a period between 10 and 24 months after a previous screening round.

To explore the influence of risk category on screening performance, the dataset was divided into risk subgroups. To compare first screening rounds with follow-up rounds, and analyze differences between subgroups, Fisher exact tests and one-way analysis of variance were used and 95% confidence intervals were calculated (23). Additionally, the difference in tumor stage, nodal stage, and grade between cancers detected in the first round and cancers detected in follow-up rounds were assessed by using χ^2 statistics. Bonferroni correction was applied for performance measures, and

a two-sided *P* value of .008 or less was considered to indicate statistical significance. All statistics were performed by using statistical software (SPSS version 22; SPSS, Chicago, Ill).

Results

A total of 577 women (mean age, 44 years ± 11; age range, 15–77 years) were recalled, some more than once (505 women were recalled once, 60 women were recalled twice, nine women were recalled three times, one woman was recalled four times, and two women were recalled five times), which resulted in 666 recalls. This led to 424 women undergoing 475 biopsies (one biopsy, 383 women; two biopsies, 35 women; three biopsies, three women; four biopsies, two women; five biopsies, one woman; mean age, 44 years ± 11; age range, 15–77 years) as a consequence of screening.

Cancers

In total, 170 newly diagnosed cancers in 156 women were documented. Of these, 129 (75.9%) cancers were screening detected and 118 cancers were detected at MR imaging (modality sensitivity, 91.5%; 47 cancers were detected at the first round and 71 cancers were detected at follow-up rounds). Mammography was available for 108 screening-detected cancers, and in 62 women a cancer was visible at mammography (modality sensitivity, 57.4%). Eleven of these cancers were found at mammography alone (all at follow-up; mean age, 52 years ± 8.8; age range, 35–69 years), of which seven cancers (63.6%) were pure ductal carcinoma in situ. The incremental cancer detection at mammography was stronger (although not significant) in women without a *BRCA* mutation than in those with a *BRCA* mutation (sensitivity increase in *BRCA* mutation carriers vs in women without a *BRCA* mutation who underwent both examinations, 3.3% [from 64.1% to 67.4%] vs 10.3% [from 75.6% to 85.9%], respectively; $P = .11$).

Forty-one interval cancers were detected outside screening, of which 16 were found because of patient symptoms. Twenty-five cancers were found at prophylactic mastectomy. Of these cancers, 21 were detected at our hospital among 421 prophylactic mastectomies in 246 patients (cancer frequency per breast, 5.0%). The frequency of prophylactic mastectomy varied among risk categories, and most were performed in carriers of the *BRCA1* mutation (265 mastectomies in *BRCA1,* 111 mastectomies in *BRCA2,* 16 mastectomies in women with a family history, 27 mastectomies in women with a personal history, and two mastectomies in the group with other reasons). Eighteen of 25 (72.0%) cancers detected at prophylactic mastectomy were pure ductal carcinoma in situ. Patient and tumor characteristics of screening-detected and interval cancers are in shown Table 2. In general, symptomatic interval cancers have a higher tumor stage, are of higher grade, and are more often invasive compared with those found at prophylactic mastectomy ($P \leq .05$).

Screening Performance

Performance metrics are presented in Table 3, with individual risk category tables in Tables E1–E7 (online). The overall screening program had a sensitivity of 75.9% when all recognized cancers were included and a sensitivity of 90.0% when prophylactic mastectomy cancers were excluded. Whereas the sensitivity of MR imaging for invasive disease was higher than that of mammography (78.7% vs 46.2%, respectively, with inclusion of all cancers [$P < .001$]; 83.3% vs 48.0%, respectively, excluding cancers detected at prophylactic mastectomy [$P < .001$]), the sensitivities of mammography and MR imaging for ductal carcinoma in situ were not significantly different (45.2% vs 41.9%, respectively, with the inclusion of all cancers [$P = .82$] and 70.0% vs 72.0%, respectively, excluding cancers detected at prophylactic mastectomy [$P = .99$]; Table 4).

The sensitivity of the screening program appears to be higher in carriers of the *BRCA2* mutation (69.7% including cancers found at prophylactic mastectomy and 92.0% excluding cancers found at prophylactic mastectomy), women with a family history positive for cancer (90.9% and 95.2% with and without cancers found at prophylactic mastectomy, respectively), women with a personal history of breast cancer (82.0% and 91.1% with and without cancers found at prophylactic mastectomy, respectively), and other patients at an increased level (100% both with and without cancers found at prophylactic mastectomy) compared with carriers of the *BRCA1* mutation (66.1% and 81.3% with and without cancers found at prophylactic mastectomy, respectively), though this did not reach statistical significance ($P = .05$ and $P = .25$ with and without cancers found at prophylactic mastectomy, respectively).

Overall recall ranged from 5.8% to 10.3% (from 57.8/1000 to 103.2/1000) across subpopulations and was substantially lower at follow-up examinations than at first-round examinations ($P < .001$; Tables E1–E7 [online]). Positive predictive value of recall was 0.19 and was lowest for women with a family history of breast cancer (positive predictive value of recall, 0.09) (Table E5 [online]). At follow-up rounds, the overall cancer detection rate was lower in all subgroups, which shows the large fraction of prevalent cancers detected at the first screening round. The largest decrease in cancer detection rate was found in carriers of the *BRCA1* mutation and other at increased breast cancer risk (a difference of 23.0 [$P = .008$] and 42.1 [$P = .05$], respectively). Nonetheless, we did not detect a difference in pT stage ($P = .12$), pN stage ($P = .63$), or tumor grade ($P = .33$) between cancers detected at the first round and follow-up rounds. Despite the declining cancer detection rate in follow-up rounds because of a lower prevalence of cancer, sensitivity of MR imaging (69.4%) was higher than that of mammography (45.9%; $P < .001$). Recall rate and biopsy rate strongly declined in follow-up rounds, whereas positive predictive value of recall and positive predictive value of biopsy increased for the entire population at both MR imaging ($P = .004$ and $P = .005$, respectively) and mammography

Table 2

Patient and Cancer Characteristics of Cancers Detected in High-Risk Screening Patients

Parameter	Screening-detected Cancers	Interval Cancers Due to Patient Symptoms	Cancers Found at Prophylactic Mastectomy	*P* Value
No. of cancers	129 (75.9)	16 (9.4)	25 (14.7)	NA
Median age at detection (y)*	50 (24–77)	42 (27–55)	42 (27–70)	<.001
Risk category				.10
BRCA1 mutation	39 (66.1)	9 (15.3)	11 (18.6)	
Untested *BRCA1* †	0 (0)	0 (0)	0 (0)	
BRCA2 mutation	23 (69.7)	2 (6.1)	8 (24.2)	
Untested *BRCA2* †	0 (0)	0 (0)	0 (0)	
Family history	20 (90.9)	1 (4.5)	1 (4.5)	
Personal history	41 (82.0)	4 (8.0)	5 (10.0)	
Other patients	6 (100)	0 (0)	0 (0)	
Invasive disease	104 (81.9)	16 (12.6)	7 (5.5)	NA
Mean pathologic size (mm) *‡	14.2 (2–70)	15.5 (5–26)	5.8 (1–11)	.24
T-stage				.54
pT1	78 (83.9)	8 (8.6)	7 (7.5)	
pT2	18 (85.7)	3 (14.3)	0 (0)	
pT3	1 (100)	0 (0)	0 (0)	
pT4D	1 (100)	0 (0)	0 (0)	
Recurrence	4 (66.7)	2 (33.3)	0 (0)	
Unknown	2 (40.0)	3 (60.0)	0 (0)	
Grade				.17
Grade 1	19 (90.5)	0 (0)	2 (9.5)	
Grade 2	31 (86.1)	4 (11.1)	1 (2.8)	
Grade 3	45 (80.4)	8 (14.3)	3 (5.4)	
Unknown	9 (64.3)	4 (28.6)	1 (7.1)	
Molecular subtype				.01
Luminal A	32 (88.9)	3 (8.3)	1 (2.8)	
Luminal B	20 (95.2)	0 (0)	1 (4.8)	
HER2 type	9 (100)	0 (0)	0 (0)	
Triple negative	18 (60.0)	8 (26.7)	4 (13.3)	
Unknown	25 (80.6)	5 (16.1)	1 (3.2)	
N-stage				.15
pN0	71 (80.7)	7 (8.0)	4 (4.5)	
pN+	24 (82.8)	5 (17.2)	0 (0)	
Unknown	9 (56.3)	4 (25.0)	3 (18.8)	
M-stage				NA
pM0	52 (82.5)	5 (7.9)	6 (9.5)	
Unknown	52 (81.3)	11 (17.2)	1 (15.6)	
In situ disease	25 (58.1)	0 (0)	18 (41.9)	NA
Grade				.02
Grade 1	1 (20.0)	0 (0)	4 (80.0)	
Grade 2	12 (52.2)	0 (0)	11 (47.8)	
Grade 3	9 (90.0)	0 (0)	1 (10.0)	
Unknown	3 (60.0)	0 (0)	2 (40.0)	
N-stage				.40
pN0	21 (61.8)	0 (0)	13 (38.2)	
pN+	0 (0)	0 (0)	1 (100)	
Unknown	4 (50.0)	0 (0)	4 (50.0)	

Note.—Unless otherwise indicated, data in parentheses are percentage. *P* value of .008 or less was considered to indicate statistical significance. NA = not applicable.

* Data in parentheses are range.

† Cancers detected in untested first-degree relatives of carriers of the *BRCA* mutation.

‡ Pathologic tumor size is known for 88 screening-detected invasive lesions, for eight interval cancers, and for five cancers detected in prophylactic mastectomy specimen.

Table 3

Performance Measures for Mammography and MR Imaging Screening in the Entire Population

Parameter	Overall	First Screening Round	Follow-up Rounds	*P* Value (Comparison of First Round vs Follow-up Rounds)
No. of screening rounds*	8818	2092	6726	…
Recall rate (per 1000)	75.5 [71.5, 82.2] (666/8818)	167.8 [152.6, 183.0] (351/2092)	46.8 [42.1, 52.0] (315/6726)	<.001†
Biopsy rate (per 1000)	53.9 [50.2, 59.4] (475/8818)	113.3 [100.2, 126.0] (237/2092)	35.4 [31.6, 40.3] (238/6726)	<.001†
PPV1	0.19 [0.16, 0.22] (129/666)	0.13 [0.10, 0.16] (47/351)	0.26 [0.21, 0.31] (82/315)	<.001†
PPV3	0.27 [0.23, 0.30] (129/475)	0.20 [0.14, 0.24] (47/237)	0.34 [0.28, 0.40] (82/238)	<.001†
CDR (per 1000)	14.6 [12.1, 17.0] (129/8818)	22.5 [15.8, 27.9] (47/2092)	12.2 [9.7, 15.0] (82/6726)	.001†
Sensitivity1 (%)	75.9 [69.5, 82.4] (129/170)	81.0 [69.6, 90.2] (47/58)	73.2 [64.7, 81.2] (82/112)	.35
Sensitivity2 (%)	90.0 [83.3, 93.7] (129/145)	88.7 [77.4, 95.6] (47/53)	89.1 [81.4, 94.7] (82/92)	.99
Specificity (%)	93.8 [93.3, 94.3] (8111/8648)	85.1 [83.4, 86.5] (1730/2034)	96.5 [96.0, 96.9] (8381/6614)	<.001†
Mammography (no. of examinations)	6245	1645	4600	
Recall rate (per 1000)	38.1 [36.3, 46.0] (238/6245)	82.1 [74.5, 101.1] (135/1645)	22.4 [19.4, 28.2] (103/4600)	<.001†
Biopsy rate (per 1000)	24.7 [22.9, 30.8] (154/6245)	46.2 [39.5, 60.0] (76/1645)	17.0 [14.6, 22.4] (78/4600)	<.001†
PPV1	0.26 [0.20, 0.31] (62/238)	0.14 [0.09, 0.20] (19/135)	0.42 [0.33, 0.52] (43/103)	<.001†
PPV3	0.40 [0.32, 0.47] (62/154)	0.25 [0.16, 0.34] (19/76)	0.55 [0.44, 0.65] (43/78)	<.001†
CDR (per 1000)	9.9 [8.1, 13.2] (62/6245)	11.6 [7.4, 18.0] (19/1645)	9.3 [7.4, 13.3] (43/4600)	.47
Sensitivity1 (%)	45.9 [35.5, 51.4] (62/135)	41.3 [27.3, 54.9] (19/46)	48.3 [35.2, 54.8] (43/89)	.47
Sensitivity2 (%)	51.7 [43.0, 60.7] (62/120)	45.2 [30.5, 59.8] (19/42)	55.1 [44.7, 66.6] (43/78)	.34
Specificity (%)	97.1 [96.7, 97.5] (5934/6110)	92.7 [91.4, 93.9] (1483/1599)	98.7 [98.3, 99.0] (4451/4511)	<.001†
MR imaging (no. of examinations)	8818	2092	6726	
Recall rate (per 1000)	61.6 [57.9, 67.7] (543/8818)	134.3 [120.2, 148.0] (281/2092)	39.0 [34.9, 44.0] (262/6726)	<.001†
Biopsy rate (per 1000)	48.1 [44.3, 53.0] (424/8818)	103.7 [90.2, 115.0] (217/2092)	30.8 [27.0, 35.2] (207/6726)	<.001†
PPV1	0.22 [0.18, 0.24] (118/543)	0.17 [0.12, 0.20] (47/281)	0.27 [0.22, 0.32] (71/262)	.004†
PPV3	0.28 [0.23, 0.31] (118/424)	0.22 [0.15, 0.26] (47/217)	0.34 [0.28, 0.40] (71/207)	.005†
CDR (per 1000)	13.4 [10.8, 15.5] (118/8818)	22.5 [15.5, 27.4] (47/2092)	10.6 [8.3, 13.1] (71/6726)	<.001†
Sensitivity1 (%)	69.4 [61.9, 75.8] (118/170)	81.0 [67.8, 89.0] (47/58)	63.4 [54.1, 72.1] (71/112)	.02
Sensitivity2 (%)	81.4 [73.7, 86.7] (118/145)	88.7 [75.3, 94.4] (47/53)	77.2 [67.5, 85.0] (71/92)	.12
Specificity (%)	95.1 [94.6, 95.5] (8223/8648)	88.5 [87.0, 89.8] (1800/2034)	97.1 [96.7, 97.5] (6423/6614)	<.001†

Note.—Data in brackets are 95% confidence intervals; data in parentheses are numerator/denominator. The *BRCA* population also included untested first degree relatives. Sensitivity1 refers to sensitivity including all cancers, sensitivity2 refers to sensitivity excluding cancers detected at prophylactic mastectomies as false-negative findings, and specificity includes cancers detected at prophylactic mastectomy as false-negative findings. CDR = cancer detection rate, PPV1 = positive predictive value of recall, PPV3 = positive predictive value of biopsy.

* The complete regimen consisted of a combination of MR imaging and mammography, when available.

† Statistically significant.

($P < .001$ for both positive predictive values) except for the subgroup of other patients at increased breast cancer risk (Table E7 [online]). For MR imaging, the availability of previous imaging examinations (MR imaging and/or mammography) strongly improved the specificity to a level that was just below the specificity of mammography (97.1% vs 98.7%; $P < .001$).

Discussion

We showed that the specificity of high-risk screening in real-life practice is

Table 4

Sensitivity for Invasive Disease and Ductal Carcinoma in Situ at Mammography and MR Imaging

Parameter	Mammography	MR Imaging	*P* Value
Invasive disease			
Sensitivity1 (%)	46.2 [36.8, 55.7] (48/104)	78.7 [70.8, 85.0] (100/127)	<.001*
Sensitivity2 (%)	48.0 [38.5, 57.7] (48/100)	83.3 [75.7, 88.9] (100/120)	<.001*
DCIS			
Sensitivity1 (%)	45.2 [29.2, 62.2] (14/31)	41.9 [28.4, 56.7] (18/43)	.82
Sensitivity2 (%)	70.0 [48.1, 85.5] (14/20)	72.0 [52.4, 85.7] (18/25)	.99

Note.—Data in brackets are 95% confidence intervals; data in parentheses are numerator/denominator. Sensitivity1 refers to sensitivity including all cancers, sensitivity2 refers to sensitivity excluding cancers detected at prophylactic mastectomies as false-negative findings. DCIS = ductal carcinoma in situ.

excellent, both for mammography and breast MR imaging, especially in follow-up rounds. This is different than in previous clinical trials that reported variable and generally much lower specificity of breast MR imaging (4,10,13,15). In follow-up rounds, the specificity of MR imaging approached the specificity of mammography, which reaffirmed that new examinations must be evaluated together with previous studies. By doing so, the number of recalls is reduced and the positive predictive value of both recall and biopsy increases.

The positive predictive value is substantially lower in women with a family history of breast cancer than in all other subgroups. This might be partly related to the much lower breast cancer risk in this population. Our results thus still indicate that strict evaluation of risk category for intensified screening is required.

In terms of sensitivity excluding cancers detected in prophylactic mastectomy, the overall performance of the screening regimen is 90.0% and varied little between subgroups, except for carriers of the *BRCA1* mutation in whom the sensitivity (81.3%) appeared to be lower. This is in line with earlier studies (24) and implies that, for carriers of the *BRCA1* mutation, an even more stringent screening regimen, or prophylactic mastectomy (25), may be indicated.

Despite the high accuracy of the screening program, it is evident that there is a reservoir of undetected cancers that are only found in prophylactic mastectomy specimens. At our center, 5% of prophylactic mastectomies contained undetected cancers. Hence, the overall sensitivity of the screening program is heavily dependent on the frequency of prophylactic mastectomies in the target population and the indication for prophylactic mastectomy. However, our results showed that the cancers found in prophylactic mastectomy specimens were in general small and usually noninvasive, which is in line with findings of previous studies (26,27). Consequently, the importance of detecting these cancers is unknown, even though they are most often detected in carriers of the *BRCA* germline mutation, which, in turn, can be partly explained by the fact that more prophylactic surgical procedures are performed in *BRCA* mutation carriers.

The addition of mammography to MR imaging increased the sensitivity of the screening program to 8.1% (from 71.9% with only MR imaging to 80.0% with both examinations) in women who underwent both examinations. These were 11 cancers, of which seven were ductal carcinoma in situ. Whether the additional detection of these cancers is beneficial to the patient is unknown. Kuhl et al (28) noted that the sensitivity of breast MR imaging for high-grade ductal carcinoma in situ is excellent, whereas mammography preferentially depicts lower grade lesions (28). In line with other studies, our results show that the sensitivity of mammography was especially low in carriers of the *BRCA* mutation and in women with a personal history of breast cancer (10,29–31). Nevertheless, the age from which supplemental mammography screening is recommended varies from country to country and may be as low as 30 years in carriers of the *BRCA* mutation (1). This advice could be modified to reflect the importance of MR screening in this population.

Our study has limitations. This was a single institutional study, which potentially limits its ability to generalize. Because the study is longitudinal in nature, clinical and imaging protocols evolved over time. Furthermore, 13 radiologists reported the cases over the years. Unfortunately, because of the limited number of cancer detections per year, the influence of changing protocols and different levels of experience cannot be assessed. Whereas we analyzed the performance of MR imaging and mammography separately, in most women these examinations were simultaneously evaluated. Therefore, the performance of each of the modalities may have been affected by the findings in the other modality. In addition, because we only have information on the prophylactic mastectomies negative for cancer that were performed at our hospital, the actual frequency of cancer in specimens from prophylactic mastectomy might be slightly over- or understated because of selection bias. Another limitation is the fact that young women who are *BRCA* mutation carriers (<30 years) are only screened with MR imaging, and hence the number of MR examinations was different from the number of mammographic examinations, which could potentially skew accuracy values. Finally, the rate of prophylactic mastectomy reduces the apparent sensitivity for all cancers, so any difference in sensitivity in breast cancer patients may largely be because of the higher rate of such mastectomies rather than an unique difficulty in detection.

From our results it is evident that the group of women who underwent intensified screening because of a family history of breast cancer was at substantially lower risk than the other subgroups. Because of the retrospective nature of the study, we were unable to obtain the actual family history and hence could not further stratify these women. Similarly, the subgroup with risk category "other" remains heterogeneous, but because of the low frequency of the specific risks within this group, further division was not possible. Finally, survival data were not available. Further studies need to establish whether there is a survival benefit for women who participate in breast cancer screening programs (32,33).

In conclusion, this study shows that in real-life practice, the performance of a high-risk screening program is affected by risk category and the frequency of prophylactic mastectomies. The specificity of the screening program is high and improves at follow-up. Especially in women who are *BRCA* mutation carriers, the added value of mammography is limited.

Acknowledgments: We thank the registration team of the Netherlands Comprehensive Cancer Organisation (IKNL) for the collection of data for the Netherlands Cancer Registry and IKNL staff for scientific advice.

Disclosures of Conflicts of Interest: S.V. disclosed no relevant relationships. **A.G.M.** disclosed no relevant relationships. **M.S.S.V.** disclosed no relevant relationships. **P.B.** disclosed no relevant relationships. **C.H.v.G.** Activities related to the present article: disclosed no relevant

relationships. Activities not related to the present article: disclosed grants and payments for lectures from Bayer Healthcare. Other relationships: disclosed no relevant relationships. **N.H.** disclosed no relevant relationships. **N.K.** Activities related to the present article: disclosed no relevant relationships. Activities not related to the present article: disclosed personal fees from QView Medical and Screenpoint Medical; disclosed shareholder in Volpara Solutions, QView Medical, and Screenpoint Medical. Other relationships: disclosed no relevant relationships. **R.M.M.** Activities related to the present article: disclosed no relevant relationships. Activities not related to the present article: disclosed consultancies from Siemens Healthineers and Screenpoint Medical; disclosed grants or grants pending from Siemens Healthineers, Seno Medical, IDS, Bayer Healthcare, and Micrima Medical. Other relationships: disclosed no relevant relationships.

References

1. Mann RM, Kuhl CK, Kinkel K, Boetes C. Breast MRI: guidelines from the European Society of Breast Imaging. Eur Radiol 2008;18(7):1307–1318.
2. Saslow D, Boetes C, Burke W, et al. American Cancer Society guidelines for breast screening with MRI as an adjunct to mammography. CA Cancer J Clin 2007;57(2):75–89. [Published correction appears in CA Cancer J Clin 2007;57(3):185.]
3. Chen S, Parmigiani G. Meta-analysis of BRCA1 and BRCA2 penetrance. J Clin Oncol 2007;25(11):1329–1333.
4. Leach MO, Boggis CR, Dixon AK, et al. Screening with magnetic resonance imaging and mammography of a UK population at high familial risk of breast cancer: a prospective multicentre cohort study (MARIBS). Lancet 2005;365(9473):1769–1778.
5. Narod SA. Screening of women at high familial risk of breast cancer. Womens Health (Lond) 2010;6(4):509–510.
6. Brekelmans CT, Seynaeve C, Bartels CC, et al. Effectiveness of breast cancer surveillance in BRCA1/2 gene mutation carriers and women with high familial risk. J Clin Oncol 2001;19(4):924–930.
7. Travis LB, Hill DA, Dores GM, et al. Breast cancer following radiotherapy and chemotherapy among young women with Hodgkin disease. JAMA 2003;290(4):465–475.
8. Degnim AC, Visscher DW, Berman HK, et al. Stratification of breast cancer risk in women with atypia: a Mayo cohort study. J Clin Oncol 2007;25(19):2671–2677.
9. Bedrosian I, Mick R, Orel SG, et al. Changes in the surgical management of patients with breast carcinoma based on preoperative magnetic resonance imaging. Cancer 2003;98(3):468–473.
10. Warner E, Plewes DB, Hill KA, et al. Surveillance of BRCA1 and BRCA2 mutation carriers with magnetic resonance imaging, ultrasound, mammography, and clinical breast examination. JAMA 2004;292(11):1317–1325.
11. Sardanelli F, Podo F, Santoro F, et al. Multicenter surveillance of women at high genetic breast cancer risk using mammography, ultrasonography, and contrast-enhanced magnetic resonance imaging (the high breast cancer risk italian 1 study): final results. Invest Radiol 2011;46(2):94–105.
12. Kuhl C, Weigel S, Schrading S, et al. Prospective multicenter cohort study to refine management recommendations for women at elevated familial risk of breast cancer: the EVA trial. J Clin Oncol 2010;28(9):1450–1457.
13. Riedl CC, Luft N, Bernhart C, et al. Triple-modality screening trial for familial breast cancer underlines the importance of magnetic resonance imaging and questions the role of mammography and ultrasound regardless of patient mutation status, age, and breast density. J Clin Oncol 2015;33(10):1128–1135.
14. Rutgers EJ, Tuut MK. CBO guideline 'Breast cancer: screening and diagnosis' [in Dutch]. Ned Tijdschr Geneeskd 2001;145(3):115–119.
15. Kriege M, Brekelmans CT, Boetes C, et al. Efficacy of MRI and mammography for breast-cancer screening in women with a familial or genetic predisposition. N Engl J Med 2004;351(5):427–437.
16. Kuhl CK, Schrading S, Leutner CC, et al. Mammography, breast ultrasound, and magnetic resonance imaging for surveillance of women at high familial risk for breast cancer. J Clin Oncol 2005;23(33):8469–8476.
17. 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(1):165–175.
18. NABON. Breast cancer guideline. CITY, the Netherlands: NABON, 2012.
19. Dalmış MU, Litjens G, Holland K, et al. Using deep learning to segment breast and fibroglandular tissue in MRI volumes. Med Phys 2017;44(2):533–546.
20. American College of Radiology. ACR BI-RADS Atlas. 5th ed. Reston, Va: American College of Radiology, 2013.
21. American College of Radiology. ACR BI-RADS Atlas. 4th ed. Reston, Va: American College of Radiology, 2003.
22. Rosenberg RD, Yankaskas BC, Abraham LA, et al. Performance benchmarks for screening mammography. Radiology 2006;241(1):55–66.
23. Newcombe RG. Two-sided confidence intervals for the single proportion: comparison of seven methods. Stat Med 1998;17(8):857–872.
24. Rijnsburger AJ, Obdeijn IM, Kaas R, et al. BRCA1-associated breast cancers present differently from BRCA2-associated and familial cases: long-term follow-up of the Dutch MRISC Screening Study. J Clin Oncol 2010;28(36):5265–5273.
25. Rebbeck TR, Friebel T, Lynch HT, et al. Bilateral prophylactic mastectomy reduces breast cancer risk in BRCA1 and BRCA2 mutation carriers: the PROSE Study Group. J Clin Oncol 2004;22(6):1055–1062.
26. Freitas V, Crystal P, Kulkarni SR, et al. The value of breast MRI in high-risk patients with newly diagnosed breast cancer to exclude invasive disease in the contralateral prophylactic mastectomy: Is there a role to choose wisely patients for sentinel node biopsy? Cancer Med 2016;5(6):1031–1036.
27. Boughey JC, Khakpour N, Meric-Bernstam F, et al. Selective use of sentinel lymph node surgery during prophylactic mastectomy. Cancer 2006;107(7):1440–1447.
28. Kuhl CK, Schrading S, Bieling HB, et al. MRI for diagnosis of pure ductal carcinoma in situ: a prospective observational study. Lancet 2007;370(9586):485–492.
29. Phi XA, Saadatmand S, De Bock GH, et al. Contribution of mammography to MRI screening in BRCA mutation carriers by BRCA status and age: individual patient data meta-analysis. Br J Cancer 2016;114(6):631–637.
30. Houssami N, Abraham LA, Miglioretti DL, et al. Accuracy and outcomes of screening mammography in women with a personal history of early-stage breast cancer. JAMA 2011;305(8):790–799.
31. Obdeijn IM, Winter-Warnars GA, Mann RM, Hooning MJ, Hunink MG, Tilanus-Linthorst MM. Should we screen BRCA1 mutation carriers only with MRI? A multicenter study. Breast Cancer Res Treat 2014;144(3):577–582.
32. Saadatmand S, Obdeijn IM, Rutgers EJ, et al. Survival benefit in women with BRCA1 mutation or familial risk in the MRI screening study (MRISC). Int J Cancer 2015;137(7):1729–1738.
33. Evans DG, Kesavan N, Lim Y, et al. MRI breast screening in high-risk women: cancer detection and survival analysis. Breast Cancer Res Treat 2014;145(3):663–672.