The NEW ENGLAND JOURNAL of MEDICINE

ESTABLISHED IN 1812 NOVEMBER 28, 2019 VOL. 381 NO. 22

Supplemental MRI Screening for Women with Extremely Dense Breast Tissue

M.F. Bakker, S.V. de Lange, R.M. Pijnappel, R.M. Mann, P.H.M. Peeters, E.M. Monninkhof, M.J. Emaus, C.E. Loo, R.H.C. Bisschops, M.B.I. Lobbes, M.D.F. de Jong, K.M. Duvivier, J. Veltman, N. Karssemeijer, H.J. de Koning, P.J. van Diest, W.P.T.M. Mali, M.A.A.J. van den Bosch, W.B. Veldhuis, and C.H. van Gils, for the DENSE Trial Study Group*

ABSTRACT

BACKGROUND

Extremely dense breast tissue is a risk factor for breast cancer and limits the detection of cancer with mammography. Data are needed on the use of supplemental magnetic resonance imaging (MRI) to improve early detection and reduce interval breast cancers in such patients.

METHODS

In this multicenter, randomized, controlled trial in the Netherlands, we assigned 40,373 women between the ages of 50 and 75 years with extremely dense breast tissue and normal results on screening mammography to a group that was invited to undergo supplemental MRI or to a group that received mammography screening only. The groups were assigned in a 1:4 ratio, with 8061 in the MRI-invitation group and 32,312 in the mammography-only group. The primary outcome was the between-group difference in the incidence of interval cancers during a 2-year screening period.

RESULTS

The interval-cancer rate was 2.5 per 1000 screenings in the MRI-invitation group and 5.0 per 1000 screenings in the mammography-only group, for a difference of 2.5 per 1000 screenings (95% confidence interval [CI], 1.0 to 3.7; P<0.001). Of the women who were invited to undergo MRI, 59% accepted the invitation. Of the 20 interval cancers that were diagnosed in the MRI-invitation group, 4 were diagnosed in the women who actually underwent MRI (0.8 per 1000 screenings) and 16 in those who did not accept the invitation (4.9 per 1000 screenings). The MRI cancer-detection rate among the women who actually underwent MRI screening was 16.5 per 1000 screenings (95% CI, 13.3 to 20.5). The positive predictive value was 17.4% (95% CI, 14.2 to 21.2) for recall for additional testing and 26.3% (95% CI, 21.7 to 31.6) for biopsy. The false positive rate was 79.8 per 1000 screenings. Among the women who underwent MRI, 0.1% had either an adverse event or a serious adverse event during or immediately after the screening.

CONCLUSIONS

The use of supplemental MRI screening in women with extremely dense breast tissue and normal results on mammography resulted in the diagnosis of significantly fewer interval cancers than mammography alone during a 2-year screening period. (Funded by the University Medical Center Utrecht and others; DENSE ClinicalTrials.gov number, NCT01315015.)

The authors' full names, academic degrees, and affiliations are listed in the Appendix. Address reprint requests to Dr. van Gils at the Julius Center for Health Sciences and Primary Care, Stratenum 6.131, University Medical Center Utrecht, P.O. Box 85500, Utrecht 3508 GA, the Netherlands, or at c.vangils@umcutrecht.nl.

*A list of members of the DENSE Trial Study Group is provided in the Supplementary Appendix, available at NEJM.org.

Drs. Bakker and de Lange and Drs. Veldhuis and van Gils contributed equally to this article.

N Engl J Med 2019;381:2091-102.
DOI: 10.1056/NEJMoa1903986
Copyright © 2019 Massachusetts Medical Society.

The New England Journal of Medicine
Downloaded from nejm.org on November 29, 2023. For personal use only. No other uses without permission.
Copyright © 2019 Massachusetts Medical Society. All rights reserved.

A Quick Take is available at NEJM.org

Women with extremely dense breast tissue have an increased risk of breast cancer, and their cancers are also less likely to be detected on mammography.[1-3] Such patients may benefit from a tailored breast-screening strategy, supplemented with more sensitive imaging methods. The benefit of supplemental imaging is the subject of a worldwide debate. In the United States, a federal law directs breast-density reporting,[4] but supplemental screening is not recommended in American guidelines.[5] Although supplemental imaging increases the rate of cancer detection in women with dense breasts,[6] the question remains whether it improves health outcomes. The first indication for a reduction in morbidity and mortality is a reduction in the incidence of interval cancers, since such a reduction may mean that cancers that would otherwise have become symptomatic would now be detected earlier.[7,8]

The Dense Tissue and Early Breast Neoplasm Screening (DENSE) trial is a randomized, controlled trial to study the effect of supplemental magnetic resonance imaging (MRI) on the incidence of interval cancers in women with extremely dense breast tissue. Here, we present the primary outcome of the first 2-year screening round of the DENSE trial.

Methods

Trial Design and Population

The trial design has been described in detail previously.[9] In the multicenter DENSE trial, we enrolled women who were participating in the Dutch population-based digital mammography screening program, which is conducted every 2 years for women between the ages of 50 and 75 years.[10,11] From December 2011 through November 2015, we enrolled screening participants who had negative results on mammography and who had extremely dense breast tissue, which was defined as grade 4 density as measured on Volpara imaging software, version 1.5 (Volpara Health Technologies).[12] Volpara density grades range from 1 to 4 (classified as "a" to "d" in the latest version) and correspond to the four-point breast-density categories of the Breast Imaging, Reporting, and Data System (BI-RADS) of the American College of Radiology, which range from almost entirely fatty tissue to extremely dense tissue. (Details regarding the grading system are provided in the Supplementary Appendix, available with the full text of this article at NEJM.org.)[12,13]

We randomly assigned the women in a 1:4 ratio to a group that was invited to undergo supplemental MRI or to a group that received mammography screening only. The women underwent central randomization with the use of a computer-based program in varying block sizes, stratified according to hospital (among the eight participating centers) and regional screening organization (among the four participating regions). After randomization, only the women who had been assigned to the MRI-invitation group were informed about the group assignment and were asked to participate in the trial. This process was performed according to the Zelen design[14] of randomization before informed consent was obtained. This design was used to prevent anxiety in the control group and to reduce the probability that women in the control group would arrange for MRI examination on their own initiative. For all women who had undergone randomization, data were gathered regarding breast density, age, socioeconomic status, and urbanization level.[15,16] The participants who underwent MRI screening received a travel allowance of €20 ($20.24 U.S.). The MRI examination was financed with grant money.

Trial Oversight

On November 11, 2011, the trial was approved by the Dutch Minister of Health, Welfare, and Sport, under advisement from the Health Council of the Netherlands. The trial was financially supported by the University Medical Center Utrecht, the Netherlands Organization for Health Research and Development, the Dutch Cancer Society, the Dutch Pink Ribbon–A Sister's Hope organization, Stichting Kankerpreventie Midden-West, and Bayer Pharmaceuticals, with an in-kind contribution from Volpara Health Technologies. The authors designed the trial, gathered and analyzed the data, and wrote the manuscript. With the exception of the University Medical Center Utrecht, none of the funders had any role in these tasks. All the authors vouch for the accuracy and completeness of the data and analysis and for the adherence of the trial to the protocol (available at NEJM.org).

Imaging

All the trial participants had a negative result on regular mammographic screening (bilateral

The New England Journal of Medicine
Downloaded from nejm.org on November 29, 2023. For personal use only. No other uses without permission.
Copyright © 2019 Massachusetts Medical Society. All rights reserved.

craniocaudal and mediolateral oblique views). A negative result was defined as a BI-RADS radiographic score of 1 or 2 on a 6-point scale on which a higher score reflects a greater cancer risk. A BI-RADS score of 6 (which indicates known biopsy-confirmed cancer) was not used for evaluation in this screening trial.[13,17] All MRI examinations were performed on 3.0 Tesla systems with the use of a dedicated bilateral breast coil. (A link to an interactive mobile app with additional data about the trial is provided in the Supplementary Appendix.)

Single-read MRI examinations were performed according to the BI-RADS MRI lexicon[18] and were conducted by breast radiologists whose experience ranged from 5 to 23 years. All the participants who had a BI-RADS score of 4 or 5 were recalled for additional workup. In participants with a BI-RADS score of 3, double reading of the MRI was performed, and if there was consensus on a score of 3, follow-up imaging with MRI after 6 months was planned. The results of the follow-up MRI had to be reported as either negative (BI-RADS score of 1 or 2, with a return to the regular screening program) or positive (BI-RADS score of 4 or 5, with recall). Women in the mammography-only group received the standard of care, which consisted of the regular screening program with invitations to undergo mammography every 2 years.

PRIMARY AND SECONDARY OUTCOMES

The primary outcome was the between-group difference in the incidence of interval cancers. We collected data regarding the number of interval cancers and the tumor characteristics in the two groups through linkage with the Netherlands Cancer Registry. In the MRI-invitation group, cancers were detected either on the screening MRI or at the 6-month repeat screening, if applicable. Interval cancers included all the breast cancers that were diagnosed after negative results on mammography before the next scheduled mammography examination. If no mammography was scheduled (e.g., because of an age of >75 years), an interval cancer was defined as one diagnosed within 24 months after the negative results on mammography. This definition presumes that the interval cancer would have been detected on subsequent mammography.

Key secondary outcomes included the recall rate for additional examination, the cancer-detection rate on MRI, the false positive rate, the positive predictive value, and tumor characteristics. The recall rate was defined as the percentage of participants who had a positive result on MRI screening among all the women who had undergone MRI screening. A BI-RADS score of 3, 4, or 5 was considered to be a positive result on MRI. For women who had more than one lesion, recall was based on the lesion with the highest BI-RADS score. The cancer-detection rate on MRI was defined as the percentage of women with a positive result on MRI screening that resulted in histologically confirmed breast cancer among all the women who had undergone MRI screening. The false positive rate was defined as the percentage of women who had a positive result on screening MRI but who were later found not to have breast cancer. In calculating the positive predictive value of recall after positive results on MRI, three measures were used. The first measure (PPV1) was the percentage of women with cancers detected on MRI screening among all the participants who had positive results on MRI. For the second measure (PPV2), the denominator consisted of all the women who had an indication for biopsy (BI-RADS score of 4 or 5). For the third measure (PPV3), the denominator consisted of all the women who had undergone biopsy.[13,17]

The program sensitivity among the women who were screened with MRI is the number of women with cancers detected on MRI screening among all the women with screening-detected or interval cancers. Descriptions were provided regarding the tumor–node–metastasis (TNM) stage, grade, morphology, and receptor status. In women with more than one tumor, we described the one with the highest TNM stage. Adverse events and serious adverse events were recorded in the trial center during or immediately after the MRI examination or reported by the women within 30 days.

STATISTICAL ANALYSIS

The trial was designed to have power of 80% to detect a between-group difference in the interval-cancer rate of 1.95 per 1000 screenings in the intention-to-screen population.[9] Interval-cancer rates were calculated as the number of interval cancers per 1000 screenings and per 1000 person-years of follow-up. Follow-up was calculated as the time from a negative result on mammogra-

The New England Journal of Medicine
Downloaded from nejm.org on November 29, 2023. For personal use only. No other uses without permission.
Copyright © 2019 Massachusetts Medical Society. All rights reserved.

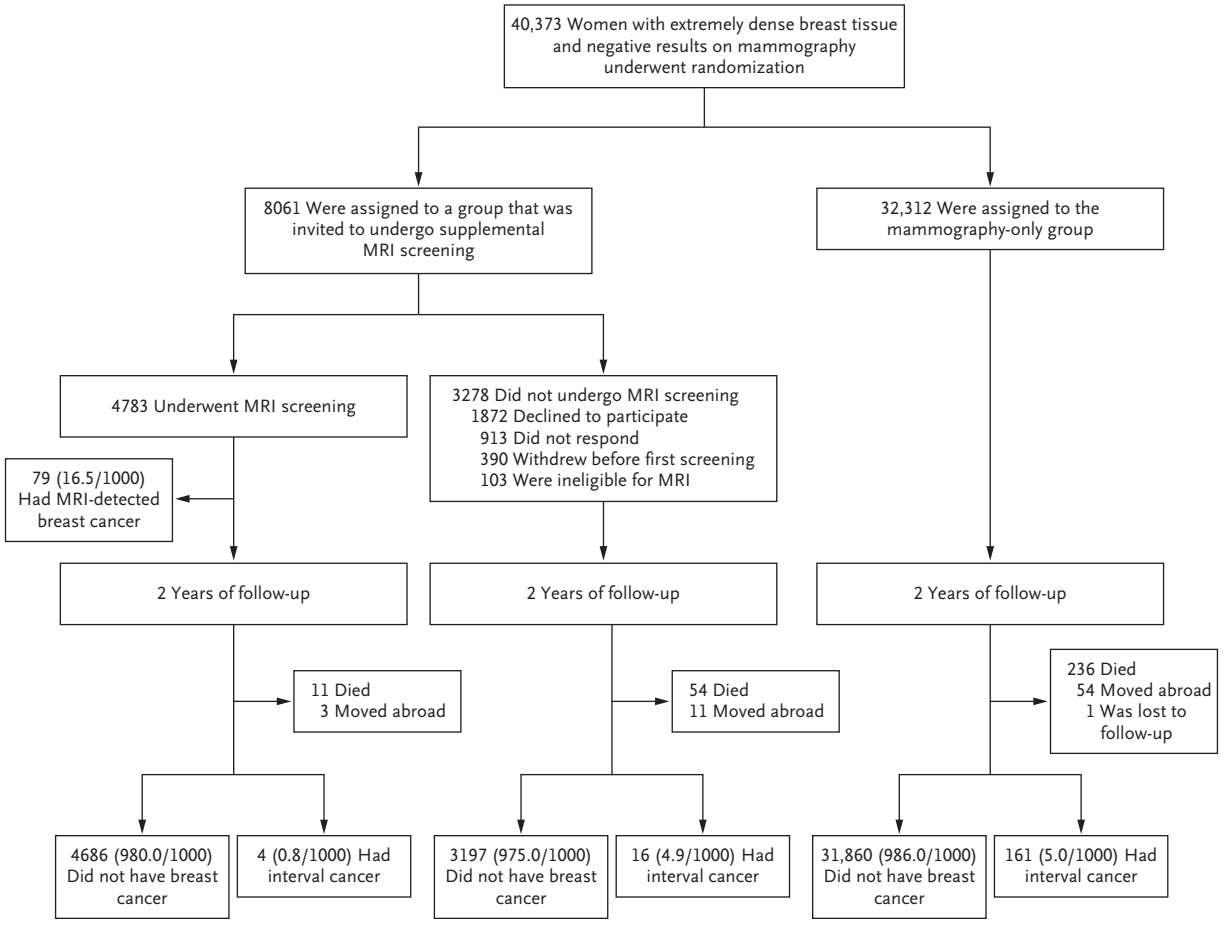

Figure 1. Enrollment and Follow-up in 2-Year Screening Program.

The women, who were between the ages of 50 and 75 years and had extremely dense breast tissue, were assigned in a 1:4 ratio to a group that was invited to undergo supplemental magnetic resonance imaging (MRI) screening or to a group that received mammography screening only. Approximately 59% of the participants in the MRI-invitation group actually underwent the supplemental procedure. Interval-cancer rates are the number of interval cancers diagnosed per 1000 screenings.

phy until a diagnosis of breast cancer, emigration, loss to follow-up, or death or until the next screening mammography took place or was scheduled to take place according to invitation, whichever occurred first. If the women were no longer invited to participate in the screening program because of age, a fixed follow-up of 24 months was adopted.

Complier average causal effect (CACE) analysis[19] was applied to estimate the effect of actually undergoing supplemental MRI screening in the subpopulation of women who said that they would have accepted MRI screening if it had been offered. CACE analysis was performed with the use of an instrumental-variables method in which the instrumental variable was the randomization to MRI invitation.[20-22] For this analysis, the interval-cancer rate among the MRI participants (i.e., those who actually underwent MRI examination) was compared with the rate among women who would have accepted MRI screening if it had been offered in the mammography-only group. We calculated 95% confidence intervals for differences in interval-cancer rates using a bootstrap-resampling method. (Details, formulas, and assumptions are provided in Fig. S2 in the Supplementary Appendix.) The type I error rate (alpha) was set at 0.05. All the analyses were performed with the use of RStudio software, version 1.0.143.

The New England Journal of Medicine
Downloaded from nejm.org on November 29, 2023. For personal use only. No other uses without permission.
Copyright © 2019 Massachusetts Medical Society. All rights reserved.

RESULTS

CHARACTERISTICS OF TRIAL POPULATION

Of the 40,373 women who underwent mammography screening, 8061 were assigned to a group that was invited to undergo MRI and 32,312 were assigned to a group that received mammography only. Of the 8061 women who were invited to undergo MRI, 4783 (59%) actually underwent MRI screening (Fig. 1 and Fig. S1). Details regarding participation have been reported previously.[16] The MRI-invitation group and the mammography-only group were well balanced with respect to baseline characteristics (Table 1).

PRIMARY OUTCOME

In the MRI-invitation group, an interval cancer was diagnosed in 20 women (4 among the MRI screening participants and 16 among the nonparticipants who were invited but did not undergo screening) of 8061. In the mammography-only group, an interval cancer was diagnosed in 161 of 32,312 women, which resulted in an interval-cancer rate of 2.5 per 1000 screenings (95% confidence interval [CI], 1.6 to 3.8) in the MRI-

Table 1. Characteristics of the Women at Baseline.*

Characteristic	MRI-Invitation Group†			Mammography-Only Group (N=32,312)
	Participants (N=4783)	Nonparticipants (N=3278)	Total (N=8061)	
Median age (IQR) — yr	54 (51–59)	56 (52–64)	55 (51–61)	54 (51–61)
Median time between mammography and MRI (IQR) — wk	10 (8–14)	NA	10 (8–14)	NA
Screening region — no. (%)‡				
Midwestern	1963 (41.0)	1365 (41.6)	3328 (41.3)	13,344 (41.3)
Eastern	1219 (25.5)	775 (23.6)	1994 (24.7)	7,992 (24.7)
Southwestern	623 (13.0)	450 (13.7)	1073 (13.3)	4,301 (13.3)
Southern	978 (20.4)	688 (21.0)	1666 (20.7)	6,674 (20.7)
Socioeconomic status — no. (%)§				
Quartile 4: highest	1828 (38.2)	1114 (34.0)	2942 (36.5)	11,646 (36.0)
Quartile 3	1144 (23.9)	775 (23.6)	1919 (23.8)	7,620 (23.6)
Quartile 2	1083 (22.6)	725 (22.1)	1808 (22.4)	7,350 (22.7)
Quartile 1: lowest	716 (15.0)	656 (20.0)	1372 (17.0)	5,655 (17.5)
Missing data	12 (0.3)	8 (0.2)	20 (0.2)	41 (0.1)
Urbanization level — no. (%)¶				
Extremely urban	871 (18.2)	703 (21.4)	1574 (19.5)	6,527 (20.2)
Strongly urban	1462 (30.6)	1080 (32.9)	2542 (31.5)	10,357 (32.1)
Moderately urban	961 (20.1)	584 (17.8)	1545 (19.2)	6,320 (19.6)
Slightly urban	714 (14.9)	467 (14.2)	1181 (14.7)	4,545 (14.1)
Not urban	713 (14.9)	400 (12.2)	1113 (13.8)	4,074 (12.6)
Missing data	62 (1.3)	44 (1.3)	106 (1.3)	489 (1.5)

* In the group that was invited to undergo magnetic resonance imaging (MRI), those who actually underwent MRI are identified as "MRI participants" and those who declined are identified as "MRI nonparticipants." Percentages may not total 100 because of rounding. IQR denotes interquartile range, and NA not applicable.

† Of the 8061 women who were invited to undergo MRI screening, 4783 (59%) actually underwent the screening.

‡ Data regarding region were missing for 1 woman in the mammography-only group.

§ The socioeconomic status according to quartile is presented as the distribution of the Dutch population in 2014. These data were available for postal codes in neighborhoods with more than 100 households.

¶ The urbanization level was determined as the number of addresses per square kilometer on the basis of postal codes. These numbers range from 0 to 499 for not urban, 500 to 999 for slightly urban, 1000 to 1499 for moderately urban, 1500 to 2499 for strongly urban, and 2500 or more for extremely urban.

The New England Journal of Medicine
Downloaded from nejm.org on November 29, 2023. For personal use only. No other uses without permission.
Copyright © 2019 Massachusetts Medical Society. All rights reserved.

Table 2. Interval-Cancer Rates and Rate Difference between Trial Groups, According to Two Analysis Methods.*

Type of Analysis	MRI-Invitation Group	Mammography-Only Group	Rate Difference (95% CI)
Intention-to-screen analysis			
Women with interval cancer — no./total no.	20/8061	161/32,312	
Interval-cancer rate (95% CI)			
No. per 1000 screenings	2.5 (1.6–3.8)	5.0 (4.3–5.8)	2.5 (1.0–3.7)
No. per 1000 person-yr	1.3 (0.7–1.8)	2.5 (2.1–2.9)	1.3 (0.6–1.9)
CACE analysis†			4.2 (2.0–6.4)
MRI participants			
Participants with interval cancer — no./total no.	4/4783	—	
Interval-cancer rate per 1000 screenings	0.8	—	
MRI nonparticipants			
Nonparticipants with interval cancer — no./total no.	16/3278	—	
Interval-cancer rate per 1000 screenings	4.9	—	
Mammography-only participants who would have accepted MRI screening if offered‡			
Women with interval cancer — no./total no.	—	97/19,172	
Interval-cancer rate per 1000 screenings	—	5.1	
Mammography-only participants who would not have accepted MRI screening if offered‡			
Women with interval cancer — no./total no.	—	64/13,140	
Interval-cancer rate per 1000 screenings	—	4.9	

* CI denotes confidence interval.
† CACE (complier average causal effect) analysis was performed with the use of an instrumental-variables method in which the instrumental variable was the randomization to MRI invitation. For this analysis, the interval-cancer rate among the MRI participants (i.e., those who actually underwent MRI examination) was compared with the rate among the women who would have accepted MRI screening if offered in the mammography-only group.
‡ The values in these analyses were not observed but were estimated on the assumption that the interval-cancer rate per 1000 screenings among the nonparticipants in the MRI-invitation group would be the same as that among potential nonparticipants in the mammography-only group.

invitation group and 5.0 per 1000 screenings (95% CI, 4.3 to 5.8) in the mammography-only group (Table 2). In the intention-to-screen analysis, the interval-cancer rate was lower by 2.5 per 1000 screenings (95% CI, 1.0 to 3.7) in the MRI-invitation group than in the mammography-only group (P<0.001).

In an analysis based on person-years, the interval-cancer rate was lower by 1.3 per 1000 person-years (95% CI, 0.6 to 1.9) in the MRI-invitation group. The exclusion of women in whom ductal carcinoma in situ (DCIS) was diagnosed did not change this result. Table S1 shows the incidence of cancers over time and includes a sensitivity analysis of different follow-up times since the receipt of negative results on MRI.

Using CACE analysis, we estimated that supplemental MRI screening among the subgroup of women who would have accepted MRI screening if it had been offered was associated with an interval-cancer rate that was lower by 4.2 per 1000 screenings (95% CI, 2.0 to 6.4) than that associated with mammography alone (P<0.001) (Fig. S2).

SECONDARY OUTCOMES

Among the 4783 MRI participants, the recall rate was 94.9 per 1000 screenings (95% CI, 86.9 to 103.6), and the cancer-detection rate with MRI was 16.5 per 1000 screenings (95% CI, 13.3 to 20.5) (Table 3). The positive predictive value of a positive MRI result (PPV1) was 17.4% (95% CI, 14.2 to 21.2), the positive predictive value of an indication for biopsy (PPV2) was 23.9% (95% CI,

The New England Journal of Medicine
Downloaded from nejm.org on November 29, 2023. For personal use only. No other uses without permission.
Copyright © 2019 Massachusetts Medical Society. All rights reserved.

Table 3. Evaluation of Supplemental MRI Screening.

Variable	Participants Who Underwent MRI Screening (N = 4783)	Rate (95% CI)
	no./total no. (%)	*no./1000 screenings*
First round of screening MRI		
Participants who were recalled for additional evaluation*	454/4783 (9.5)	
BI-RADS 3	150/454 (33.0)	
BI-RADS 4	286/454 (63.0)	
BI-RADS 5	18/454 (4.0)	
Participants who had indication for biopsy	331/4783 (6.9)	
BI-RADS 4 or 5 on first MRI	304/331 (91.8)	
BI-RADS 4 or 5 on 6-mo follow-up MRI after initial BI-RADS 3†	27/331 (8.2)	
Participants who underwent biopsy‡	300/4783 (6.3)	
After BI-RADS 4 or 5 on first MRI	276/300 (92.0)	
After BI-RADS 4 or 5 on 6-mo follow-up MRI after initial BI-RADS 3	24/300 (8.0)	
Women with confirmed cancers after positive MRI screening§	79/4783 (1.7)	
Type of cancer		
Ductal carcinoma in situ	15/79 (19.0)	
Invasive cancer	64/79 (81.0)	
Recall rate	454/4783 (9.5)	94.9 (86.9–103.6)
Biopsy rate	300/4783 (6.3)	62.7 (56.2–70.0)
All cancers		
Cancer-detection rate	79/4783 (1.7)	16.5 (13.3–20.5)
False positive rate	375/4700 (8.0)	79.8 (72.4– 87.9)
Measure of positive predictive value¶		
1	79/454 (17.4)	
2	79/331 (23.9)	
3	79/300 (26.3)	
Invasive cancers		
Cancer-detection rate	64/4783 (1.3)	13.4 (10.5–17.1)
False positive rate	390/4715 (8.3)	82.7 (75.2–90.9)
Measure of positive predictive value¶		
1	64/454 (14.1)	
2	64/331 (19.3)	
3	64/300 (21.3)	

* The assessment categories of the Breast Imaging, Reporting, and Data System (BI-RADS) of the American College of Radiology include scores ranging from 0 to 6 as follows: incomplete examination, 0; negative, 1; benign, 2; probably benign, 3; suspicious, 4; highly suggestive of cancer, 5; and known biopsy-confirmed cancer, 6. Category 6 was not used in the evaluation of screening results in this trial.

† A BI-RADS score of 4 or 5 on the follow-up MRI was the indication for biopsy.

‡ Of the 331 participants who had an indication for biopsy, 31 did not undergo the procedure because the lesion was no longer visible on additional imaging (in 18 participants), the lesion was an intramammary lymph node or cyst (in 8), biopsy of the lesion was technically not possible (in 2), or the lesion was already histologically proven to be benign (in 3).

§ Two synchronous cancers were diagnosed in 1 participant; only the tumor with the highest stage was used in the analyses. In 4 participants, breast cancer was diagnosed on biopsy after the 6-month follow-up MRI when the women had an initial BI-RADS score of 3.

¶ The positive predictive value is the proportion of participants who had confirmed breast cancer after positive results on MRI screening. Measure 1 included those who had a positive MRI result (BI-RADS score of 3, 4, or 5); measure 2, those who had an indication for biopsy (BI-RADS score of 4 or 5, including those with a BI-RADS score of 3 on initial MRI and a score of 4 or 5 on follow-up MRI); and measure 3, those who underwent biopsy (BI-RADS score of 4 or 5, including those with a score of 4 or 5 on follow-up MRI).

The New England Journal of Medicine
Downloaded from nejm.org on November 29, 2023. For personal use only. No other uses without permission.
Copyright © 2019 Massachusetts Medical Society. All rights reserved.

Table 4. Prognostic Characteristics of Cancers Detected on MRI Screening and Interval Cancers.*

Characteristic	MRI Participants (N=4783)				MRI Nonparticipants (N=3278)		Mammography-Only Group (N=32,312)	
	Cancers Detected on MRI Screening		Interval Cancers		Interval Cancers		Interval Cancers	
	no.	*no./1000 screenings*	*no.*	*no./1000 screenings*	*no.*	*no./1000 screenings*	*no.*	*no./1000 screenings*
Women with diagnosed cancer†	79		4		16		161	
Histologic type								
DCIS‡	15	3.1	0		2	0.6	9	0.3
Invasive ductal carcinoma	35	7.3	2	0.4	10	3.1	113	3.5
Invasive lobular carcinoma	9	1.9	2	0.4	4	1.2	20	0.6
Mixed invasive ductal and lobular carcinoma	8	1.7	0		0		3	0.1
Tubular carcinoma	7	1.5	0		0		2	0.1
Other invasive carcinoma	5	1.0	0		0		14	0.4
Status for axillary lymph nodes §								
Negative	70	14.6	2	0.4	9	2.7	89	2.6
Positive	9	1.9	2	0.4	7	2.1	72	2.2
Tumor stage¶‖								
Early (0 or I)	72	15.1	2	0.4	8	2.4	67	2.1
Late (II, III, or IV)	7	1.5	2	0.4	8	2.4	94	2.9
Tumor grade								
DCIS								
I, well-differentiated	6	1.3	0		0		3	0.1
II, moderately differentiated	6	1.3	0		1	0.3	1	<0.1
III, poorly differentiated	3	0.6	0		1	0.3	4	0.1
Missing data or could not be assessed	0		0		0		1	
Invasive								
I, well differentiated	31	6.5	0		0		29	0.9
II, moderately differentiated	24	5.0	2	0.4	7	2.1	70	2.2
III, poorly differentiated	4	0.8	1	0.2	4	1.2	39	1.2
Missing data or could not be assessed	5		1		3		14	

Downloaded from nejm.org on November 29, 2023. For personal use only. No other uses without permission.
The New England Journal of Medicine
Copyright © 2019 Massachusetts Medical Society. All rights reserved.

Receptor status**									
Positive for estrogen receptor, progesterone receptor, or both	56	11.7	3	0.6	10	3.1	119	3.7	
HER2 enriched	2	0.4	1	0.2	2	0.6	15	0.5	
Triple negative	4	0.8	0		1	0.3	16	0.5	
Missing data	2		0		1		2		

* DCIS denotes ductal carcinoma in situ, and HER2 human epidermal growth factor receptor 2.
† If women were found to have two synchronous breast cancers, only the tumor with the highest stage was used for all the analyses.
‡ DCIS also includes comedo type, intraductal papillary carcinoma with DCIS, and Paget's disease of the nipple (noninvasive with or without DCIS).
§ For women who underwent neoadjuvant chemotherapy, lymph-node status was determined on the basis of the clinical node stage.
¶ Tumor stage is based on the pathological tumor–node–metastasis (TNM) classification for women who did not receive neoadjuvant therapy. The clinical TNM classification is presented for women with invasive cancer who received neoadjuvant therapy and for those with no available data regarding the pathological TNM stage.
‖ With respect to invasive tumors only, the median tumor size was 9.5 mm (interquartile range, 6.8 to 12.0) among MRI participants with screening-detected tumors, 13.0 mm (interquartile range, 10.5 to 17.0) among MRI participants with interval cancers, 15.0 mm (interquartile range, 12.0 to 31.0) among MRI nonparticipants, and 17.0 mm (interquartile range, 12.0 to 23.0) among women in the mammography-only group. Data are not provided for women who underwent neoadjuvant chemotherapy (5 MRI participants, 4 MRI nonparticipants, and 30 women in the mammography-only group).
** Data regarding receptor status are provided only for women with invasive cancers.

19.6 to 28.8), and the positive predictive value of a biopsy (PPV3) was 26.3% (95% CI, 21.7 to 31.6).

The false positive rate was 79.8 per 1000 screenings (95% CI, 72.4 to 87.9) (specificity, 92%). As a result of the MRI screening, 300 women underwent a breast biopsy; of these women, breast cancer was diagnosed in 79 (64 with invasive breast cancer and 15 with DCIS).

The program sensitivity of MRI screening was 95.2% (95% CI, 88.1 to 98.7). Table 4 shows the characteristics of all the cancers detected on MRI screening and interval tumors. The screening-detected tumors were smaller on average than those in the other groups. The median size of invasive tumors was 9.5 mm (interquartile range, 6.8 to 12.0) among MRI participants with screening-detected tumors, 13.0 mm (interquartile range, 10.5 to 17.0) among MRI participants with interval cancers, 15.0 mm (interquartile range, 12.0 to 31.0) among MRI nonparticipants, and 17.0 mm (interquartile range, 12.0 to 23.0) among women in the mammography-only group.

Among the MRI participants, the absolute incidence of invasive (ductal and lobular) cancers was higher than that in the mammography-only group, as was the absolute incidence of DCIS and tubular cancers; the latter may have an indolent disease course. The absolute incidence of node-negative and early-stage cancers was also higher among MRI participants. This finding was accompanied by a slightly lower rate of late-stage cancers, but the number of such cancers that were detected was small, and a decrease in the number of late-stage cancers was not expected until after several years of follow-up.[23,24] Cancers that were detected in the MRI-invitation group appeared to be better differentiated and more often were hormone-receptor positive than those in the mammography-only group. At the next mammography screening, the cancer-detection rate was 2.0 per 1000 mammography screenings among the MRI participants, as compared with 7.1 per 1000 screenings among the MRI nonparticipants and 6.0 per 1000 screenings among the women in the mammography-only group; these findings indicate that MRI examination advanced the time of diagnosis (Table S1).

Among the MRI participants, 0.1% reported either an adverse event or a serious adverse event during or immediately after the MRI examination (Table S2). These events were related to vasovagal responses, contrast reactions, or intravenous

The New England Journal of Medicine
Downloaded from nejm.org on November 29, 2023. For personal use only. No other uses without permission.
Copyright © 2019 Massachusetts Medical Society. All rights reserved.

line infiltration. In addition, Table S2 shows all adverse events and serious adverse events that were reported by the MRI participants on a 30-day questionnaire that surveyed all health problems, regardless of a connection to the MRI examination.

DISCUSSION

In the first round of MRI screening of women between the ages of 50 and 75 years with extremely dense breast tissue (defined as Volpara grade 4 [or d]) and negative results on mammography, we observed a significantly lower interval-cancer rate than in the mammography-only group (2.5 vs. 5.0 per 1000 screenings) in the intention-to-screen analysis. Among the women who were invited to undergo MRI, 59% actually underwent the procedure. Of the 20 interval cancers diagnosed in the MRI-invitation group, 4 were diagnosed in the women who had undergone MRI and 16 in those who had not. Among the women who would have accepted MRI screening if it had been offered, the incidence of interval cancers after the use of supplemental MRI screening was estimated to be lower by 4.2 per 1000 screenings than the incidence after mammography alone on CACE analysis, which resulted in an interval-cancer rate similar to that observed on mammography in women with very fatty breasts (Volpara grade 1 [or a]).[3] Depending on the proportion of women who would accept MRI screening in clinical practice, the effect at the population level could be closer to the effect in either the intention-to-screen analysis or the CACE analysis.

Undergoing supplemental MRI was associated with a cancer-detection rate of 16.5 per 1000 screenings and resulted in a false positive rate of 8.0% (79.8 per 1000 screenings). Of the women who underwent a breast biopsy on the basis of an MRI indication, 26.3% had breast cancer and 73.7% did not.

In J-START (Japan Strategic Anti-cancer Randomized Trial),[25] in which investigators evaluated supplemental ultrasonographic breast screening among Japanese women between the ages of 40 and 49 years, 58% of the participants had heterogeneously or extremely dense breasts (similar to Volpara grade 3 or 4 [c or d]). Cancer-detection rates were 3.3 per 1000 screenings for mammography alone and 5.0 per 1000 screenings for mammography plus ultrasonography. The addition of ultrasonographic screening resulted in an interval-cancer rate of 0.5 per 1000 screenings, as compared with 1.0 per 1000 screenings with mammography alone, and an increase in the false positive rate from 8.8% to 12.6%. In the Japanese trial, the baseline interval-cancer rates were much lower than those in our trial, which may be related to a lower risk of breast cancer in this population; other factors include the 1-year interval between screenings and the lack of preselection of women with extremely dense breasts.

In several paired studies of MRI screening and mammography involving women with dense breast tissue, all the women underwent MRI, so investigators could not evaluate the effect of such screening on the interval-cancer rate. In a study involving 612 women between the ages of 25 and 91 years who had breast density similar to Volpara grade 3 or 4 (c or d) and at least one other risk factor for breast cancer, Berg et al.[26] observed a cancer-detection rate with MRI and mammography together that was higher by 18 per 1000 screenings than the rate with mammography alone; this higher detection rate was accompanied by an increase of more than 20% in the incidence of false positive results. In a trial involving 478 Chinese women between the ages of 30 and 71 years who had dense breast tissue and normal results on mammography, Chen et al.[27] observed a cancer-detection rate of 33 per 1000 screenings with MRI, which was accompanied by a false positive rate of 5.2% (with only BI-RADS scores of 4 or 5 considered to be positive). The indication for the breast examination (e.g., screening or evaluation because of symptoms) in this study is not entirely clear. In a trial involving 2120 women between the ages of 40 and 70 years that included 60% who had dense breast tissue (and approximately 20% with extremely dense breast tissue), Kuhl et al.[28] found that a first round of MRI in participants with normal findings on mammography, ultrasonography, or both resulted in a cancer-detection rate of 22.6 per 1000 screenings and was accompanied by a false positive rate of 9.7% (with a BI-RADS score of 3, 4, or 5 considered to be a positive outcome). In our trial, a cancer-detection rate of 16.5 per 1000 screenings for the first round of MRI screening and a false positive rate of 8.0% appear to be roughly in line with these findings. A direct comparison of exact numbers is difficult because of differences in populations but also because of

The New England Journal of Medicine
Downloaded from nejm.org on November 29, 2023. For personal use only. No other uses without permission.
Copyright © 2019 Massachusetts Medical Society. All rights reserved.

the limited sample size of most of the other studies, which creates wide confidence intervals around estimates.

The main strength of our trial is the randomized design, which allowed us to study the effect of MRI screening on the interval-cancer rate as the primary outcome. Other strengths of our trial are its embedding in a population-based screening program, the multicenter design with different centers using standardized MRI protocols from different vendors, the use of fully automated and volumetric measurements of mammographic density, and the completeness of the data collection. The women were selected solely because they had extremely dense breast tissue and not because of other risk factors. In the upcoming years, the two incident screening rounds will provide important information on the value of ongoing supplemental MRI screening as compared with a one-time-only supplemental screening. In the prevalent screening round described here, we may have detected an increased number of slow-growing cancers that were less aggressive and that had been present for a long time. This hypothesis is also indicated by the relatively large number of well-differentiated and hormone-receptor–positive cancers among the MRI participants. It is unclear how many of the cancers detected in our trial were life-threatening and what fraction, if any, represents overdiagnosis.

A limitation of our trial is that it is not large enough to look at the effect of MRI screening on breast cancer–specific or overall mortality. This outcome would require a much larger sample size and longer follow-up. The lower rate of interval cancers that we found among participants who underwent MRI is indicative of and prerequisite for an effect on mortality.[7] After that, a reduction in the number of advanced cancers would also be required to show a mortality benefit, which would require several years of follow-up.[23,24] Thus, we are now using our results in a simulation study to evaluate the reduction in mortality and the extent of overdiagnosis, together with the effects on costs and quality of life.[29-31] The recall rate of 94.9 per 1000 screenings on MRI is a concern for potential implementation of supplemental screening. Therefore, we are now evaluating methods for minimizing false positive outcomes (e.g., by computer-aided diagnosis, radiomics, and deep-learning methods). The issue of reducing the costs of MRI screening will be addressed by validating the use of abbreviated MRI protocols[28] in this population.

We found that supplemental screening with MRI in women with extremely dense breast tissue resulted in the diagnosis of significantly fewer interval cancers than the use of mammography alone. The data from incident screening rounds and longer follow-up are needed in combination with simulation studies to assess the effect on the rate of advanced cancers and, eventually, on mortality.

A data sharing statement provided by the authors is available with the full text of this article at NEJM.org.

Presented at the Clinical Trials in Radiology session at the European Congress of Radiology 2019.

Supported by the University Medical Center Utrecht (project number, UMCU DENSE), the Netherlands Organization for Health Research and Development (project numbers, ZonMW-200320002-UMCU and ZonMW Preventie 50-53125-98-014), the Dutch Cancer Society (project numbers, DCS-UU-2009-4348, UU-2014-6859, and UU-2014-7151), the Dutch Pink Ribbon–A Sister's Hope (project number, Pink Ribbon-10074), Bayer Pharmaceuticals (project number, BSP-DENSE), and Stichting Kankerpreventie Midden-West. For research purposes, Volpara Health Technologies provided Volpara Imaging Software, version 1.5, for installation on servers in the screening units.

Disclosure forms provided by the authors are available with the full text of this article at NEJM.org.

We thank the trial participants for their contributions; the regional screening organizations, Volpara Health Technologies, the Dutch Expert Center for Screening, and the National Institute for Public Health and the Environment for their advice and in-kind contributions; the data managers and project support personnel for their assistance during the trial; and R.K. Stellato, M.Sc., for her assistance with the statistical analyses and F.J. Gilbert, M.D., Ph.D., for her critical review of a previous version of the manuscript.

APPENDIX

The authors' full names and academic degrees are as follows: Marije F. Bakker, Ph.D., Stéphanie V. de Lange, M.D., Ruud M. Pijnappel, M.D., Ph.D., Ritse M. Mann, M.D., Ph.D., Petra H.M. Peeters, M.D., Ph.D., Evelyn M. Monninkhof, Ph.D., Marleen J. Emaus, Ph.D., Claudette E. Loo, M.D., Ph.D., Robertus H.C. Bisschops, M.D., Ph.D., Marc B.I. Lobbes, M.D., Ph.D., Matthijn D.F. de Jong, M.D., Katya M. Duvivier, M.D., Jeroen Veltman, M.D., Ph.D., Nico Karssemeijer, Ph.D., Harry J. de Koning, M.D., Ph.D., Paul J. van Diest, M.D., Ph.D., Willem P.T.M. Mali, M.D., Ph.D., Maurice A.A.J. van den Bosch, M.D., Ph.D., Wouter B. Veldhuis, M.D., Ph.D., and Carla H. van Gils, Ph.D.

The authors' affiliations are as follows: the Julius Center for Health Sciences and Primary Care (M.F.B., S.V.L., P.H.M.P., E.M.M., C.H.G.) and the Departments of Radiology (S.V.L., R.M.P., M.J.E., W.P.T.M.M., M.A.A.J.B., W.B.V.) and Pathology (P.J.D.), University Medical Center Utrecht, Utrecht University, Utrecht, the Dutch Expert Center for Screening (R.M.P.) and the Department of Radiology, Radboud University Nijmegen Medical Center (R.M.M., N.K.), Nijmegen, the Department of Radiology, Antoni van Leeuwenhoek

The New England Journal of Medicine
Downloaded from nejm.org on November 29, 2023. For personal use only. No other uses without permission.
Copyright © 2019 Massachusetts Medical Society. All rights reserved.

Hospital (C.E.L.), and the Department of Radiology and Nuclear Medicine, Cancer Center Amsterdam, Amsterdam UMC, Vrije Universiteit Amsterdam (K.M.D.), Amsterdam, the Department of Radiology, Albert Schweitzer Hospital, Dordrecht (R.H.C.B.), the Department of Radiology and Nuclear Medicine, Maastricht University Medical Center, Maastricht, and the Department of Medical Imaging, Zuyderland Medical Center, Sittard-Geleen (M.B.I.L.), the Department of Radiology, Jeroen Bosch Hospital, 's-Hertogenbosch (M.D.F.J.), the Department of Radiology, Hospital Group Twente, Almelo (J.V.), and the Department of Public Health, Erasmus Medical Center, Rotterdam (H.J.K.) — all in the Netherlands; and the Department of Epidemiology and Biostatistics, School of Public Health, Faculty of Medicine, Imperial College London, London (P.H.M.P.).

REFERENCES

1. Boyd NF, Guo H, Martin LJ, et al. Mammographic density and the risk and detection of breast cancer. N Engl J Med 2007;356:227-36.
2. Mandelson MT, Oestreicher N, Porter PL, et al. Breast density as a predictor of mammographic detection: comparison of interval- and screen-detected cancers. J Natl Cancer Inst 2000;92:1081-7.
3. Wanders JO, Holland K, Veldhuis WB, et al. Volumetric breast density affects performance of digital screening mammography. Breast Cancer Res Treat 2017;162:95-103.
4. DenseBreast-info. Is there a national reporting standard? March 28, 2019 (https://densebreast-info.org/is-there-a-federal-law.aspx).
5. Siu AL. Screening for breast cancer: U.S. Preventive Services Task Force recommendation statement. Ann Intern Med 2016;164:279-96.
6. Melnikow J, Fenton JJ, Whitlock EP, et al. Supplemental screening for breast cancer in women with dense breasts: a systematic review for the U.S. Preventive Services Task Force. Ann Intern Med 2016;164:268-78.
7. Irwig L, Houssami N, Armstrong B, Glasziou P. Evaluating new screening tests for breast cancer. BMJ 2006;332:678-9.
8. Lord SJ, Irwig L, Simes RJ. When is measuring sensitivity and specificity sufficient to evaluate a diagnostic test, and when do we need randomized trials? Ann Intern Med 2006;144:850-5.
9. Emaus MJ, Bakker MF, Peeters PH, et al. MR imaging as an additional screening modality for the detection of breast cancer in women aged 50-75 years with extremely dense breasts: the DENSE trial study design. Radiology 2015;277:527-37.
10. National evaluation of breast cancer screening in the Netherlands. Rotterdam, the Netherlands: National Institute for Public Health and the Environment, 2014.
11. Timmers JM, den Heeten GJ, Adang EM, Otten JD, Verbeek AL, Broeders MJ. Dutch digital breast cancer screening: implications for breast cancer care. Eur J Public Health 2012;22:925-9.
12. van Engeland S, Snoeren PR, Huisman H, Boetes C, Karssemeijer N. Volumetric breast density estimation from full-field digital mammograms. IEEE Trans Med Imaging 2006;25:273-82.
13. D'Orsi CJ, Sickles EA, Mendelson EB, Morris EA. ACR BI-RADS atlas: breast imaging reporting and data system. 5th ed. Reston, VA: American College of Radiology, 2013.
14. Zelen M. A new design for randomized clinical trials. N Engl J Med 1979;300:1242-5.
15. Kerncijfers wijken en buurten. Statistics Netherlands, 2014 (http://statline.cbs.nl/Statweb/).
16. de Lange SV, Bakker MF, Monninkhof EM, et al. Reasons for (non)participation in supplemental population-based MRI breast screening for women with extremely dense breasts. Clin Radiol 2018;73(8):759.e1-759.e9.
17. D'Orsi CJ, Mendelson EB, Ikeda DM. Breast Imaging Reporting and Data System: ACR BI-RADS breast imaging atlas. Reston, VA: American College of Radiology, 2003.
18. Ikeda DM, Hylton NM, Kuhl CK. BI-RADS: magnetic resonance imaging. Reston, VA: American College of Radiology, 2003.
19. Cuzick J, Edwards R, Segnan N. Adjusting for non-compliance and contamination in randomized clinical trials. Stat Med 1997;16:1017-29.
20. Angrist JD, Imbens GW, Rubin DB. Identification of causal effects using instrumental variables. J Am Stat Assoc 1996;91:444-55.
21. Dunn G, Maracy M, Tomenson B. Estimating treatment effects from randomized clinical trials with noncompliance and loss to follow-up: the role of instrumental variable methods. Stat Methods Med Res 2005;14:369-95.
22. Stuart EA, Perry DF, Le HN, Ialongo NS. Estimating intervention effects of prevention programs: accounting for noncompliance. Prev Sci 2008;9:288-98.
23. Foca F, Mancini S, Bucchi L, et al. Decreasing incidence of late-stage breast cancer after the introduction of organized mammography screening in Italy. Cancer 2013;119:2022-8.
24. Tabár L, Yen AM, Wu WY, et al. Insights from the breast cancer screening trials: how screening affects the natural history of breast cancer and implications for evaluating service screening programs. Breast J 2015;21:13-20.
25. Ohuchi N, Suzuki A, Sobue T, et al. Sensitivity and specificity of mammography and adjunctive ultrasonography to screen for breast cancer in the Japan Strategic Anti-cancer Randomized Trial (J-START): a randomised controlled trial. Lancet 2016;387:341-8.
26. Berg WA, Zhang Z, Lehrer D, et al. Detection of breast cancer with addition of annual screening ultrasound or a single screening MRI to mammography in women with elevated breast cancer risk. JAMA 2012;307:1394-404.
27. Chen SQ, Huang M, Shen YY, Liu CL, Xu CX. Application of abbreviated protocol of magnetic resonance imaging for breast cancer screening in dense breast tissue. Acad Radiol 2017;24:316-20.
28. Kuhl CK, Strobel K, Bieling H, et al. Supplemental breast MR imaging screening of women with average risk of breast cancer. Radiology 2017;283:361-70.
29. Habbema JD, van Oortmarssen GJ, Lubbe JT, van der Maas PJ. The MISCAN simulation program for the evaluation of screening for disease. Comput Methods Programs Biomed 1985;20:79-93.
30. Mandelblatt JS, Stout NK, Schechter CB, et al. Collaborative modeling of the benefits and harms associated with different U.S. breast cancer screening strategies. Ann Intern Med 2016;164:215-25.
31. Sankatsing VD, Heijnsdijk EA, van Luijt PA, van Ravesteyn NT, Fracheboud J, de Koning HJ. Cost-effectiveness of digital mammography screening before the age of 50 in the Netherlands. Int J Cancer 2015;137:1990-9.

Copyright © 2019 Massachusetts Medical Society.

The New England Journal of Medicine
Downloaded from nejm.org on November 29, 2023. For personal use only. No other uses without permission.
Copyright © 2019 Massachusetts Medical Society. All rights reserved.