Received: 25 November 2022 | Revised: 10 March 2023 | Accepted: 21 March 2023

DOI: 10.1002/ijc.34530

CANCER EPIDEMIOLOGY

IJC INTERNATIONAL JOURNAL of CANCER

Cumulative risks of false positive recall and screen-detected breast cancer after multiple screening examinations

Lindy M. Kregting[1] | Nicolien T. van Ravesteyn[1] | Sarocha Chootipongchaivat[1] | Eveline A. M. Heijnsdijk[1] | Johannes D. M. Otten[2] | Mireille J. M. Broeders[2,3] | Harry J. de Koning[1]

[1]Department of Public Health, Erasmus MC, University Medical Center Rotterdam, Rotterdam, The Netherlands

[2]Department for Health Evidence, Radboudumc, Nijmegen, The Netherlands

[3]Dutch Expert Centre for Screening, Nijmegen, The Netherlands

Correspondence
Lindy M. Kregting, Erasmus MC, Department of Public Health, Dr. Molewaterplein 40, 3015 GD Rotterdam, The Netherlands.
Email: l.kregting@erasmusmc.nl

Funding information
Rijksinstituut voor Volksgezondheid en Milieu

Abstract

Women tend to make a decision about participation in breast cancer screening and adhere to this for future invitations. Therefore, our study aimed to provide high-quality information on cumulative risks of false-positive (FP) recall and screen-detected breast cancer over multiple screening examinations. Individual Dutch screening registry data (2005-2018) were gathered on subsequent screening examinations of 92 902 women age 49 to 51 years in 2005. Survival analyses were used to calculate cumulative risks of a FP and a true-positive (TP) result after seven examinations. Data from 66 472 women age 58 to 59 years were used to extrapolate to 11 examinations. Participation, detection and additional FP rates were calculated for women who previously received FP results compared to women with true negative (TN) results. After 7 examinations, the cumulative risk of a TP result was 3.7% and the cumulative risk of a FP result was 9.1%. After 11 examinations, this increased to 7.1% and 13.5%, respectively. Following a FP result, participation was lower (71%-81%) than following a TN result (>90%). In women with a FP result, more TP results (factor 1.59 [95% CI: 1.44-1.72]), more interval cancers (factor 1.66 [95% CI: 1.41-1.91]) and more FP results (factor 1.96 [95% CI: 1.87-2.05]) were found than in women with TN results. In conclusion, due to a low recall rate in the Netherlands, the cumulative risk of a FP recall is relatively low, while the cumulative risk of a TP result is comparable. Breast cancer diagnoses and FP results were more common in women with FP results than in women with TN results, while participation was lower.

KEYWORDS
breast neoplasms, false positive recall, mass screening, screen detection

What's new?

Population-based breast cancer screening programmes reduce breast cancer mortality. However, presenting the potential risks over multiple screening examinations is crucial to enable

Abbreviations: BI-RADS, Breast Imaging-Reporting and Data System; DBT, digital breast tomosynthesis; FN, False negative; FNDA, false negative diagnostic assessment; FP, false positive; HR, hazard rate; IKNL, The Netherlands Comprehensive Cancer Organisation; NKR, Dutch Cancer Registry; RR, relative risk; TN, true negative; TP, true positive; US, the United States.

women to make an informed choice about participation. In this breast cancer screening nationwide registry study using 13 years of follow-up data from the Netherlands, the cumulative risk of a false-positive recall was relatively low, while the cumulative risk of a true-positive result was comparable to that in other European countries. The rates of screen-detected and interval cancers and false-positives were higher in women who had received false-positive results than in women with true-negative results, while their participation was lower.

1 | INTRODUCTION

Population-based breast cancer screening programmes have been shown to reduce breast cancer mortality by detecting breast cancers earlier.[1] Because of this, many Western countries have implemented a national or regional breast cancer screening programme for their citizens.[2] Within these programmes, women between age 50 and 69, but sometimes also slightly younger or older, are invited for breast cancer screening annually, biennially or triennially.[2] This means that these women are invited to participate in multiple breast cancer screening examinations during their life.

In addition to the reduction in breast cancer mortality, the detection of earlier stage cancers also leads to less invasive treatment, potentially leading to an increase in quality of life.[3] However, breast cancer screening is also associated with harms including overdiagnosis and false positive (FP) screening results.[4,5] Many studies investigated the extent of these harms to be able to weight them against the benefits, but also to be able to inform the invited women so they can make an informed decision whether to participate or not. In these studies the extent of overdiagnoses and FP results of a screening programme were found to differ substantially between countries.[6,7] These differences in FP rate can mainly be attributed to differing aspects of the screening programmes, such as programme organisation (ie, extent of centralisation, single vs double reading, experience of radiologists, screening interval and age of the population invited) and cultural factors (ie, risk aversion and litigation culture).[6,7] For example, the specificity of subsequent breast cancer screening examinations in Denmark is considerably higher than that in the United States (US; 99% compared to 92%) and Denmark, thus, has a substantially lower FP rate.[7]

In the Netherlands, women are invited for breast cancer screening with digital mammography biennially between the ages of 50 and 74. The programme has a relatively low recall rate of 2.4% which leads to a FP rate of 1.7%.[8] This percentage also represents the average risk of a screening test resulting in a FP result. However, since women are invited up to 13 times in their lives, it is important to provide high-quality information on the cumulative risks over multiple screening examinations to enable women to make an informed decision about participating. Re-attendance is high in the Netherlands and it is suggested that most women make a decision about participation and adhere to this decision for future invitations.[8,9] Therefore, presenting risks over multiple screening examinations is crucial to enable women to make an informed choice. Furthermore, the rate of FP results per true positive (TP) result gives an indication of the balance between short-term screening benefits and harms in a specific screening programme. It is known that this rate is higher in the initial screening examination than in subsequent examinations.[8,10] However, it is uncertain what this rate will be over multiple examinations cumulatively.

Several studies have analysed cumulative risks of FP results over multiple screening examinations in different countries and found ranges from 8% to 61% over 10 examinations for women with an average breast cancer risk.[11-16] The biggest difference was seen when comparing results from studies in the United States to those in Europe, due to the difference in screening interval and recall rate. Within Europe, where screening intervals and recall rates are more comparable, only a few countries calculated cumulative risks. Despite this comparability in programme, the cumulative risks still ranged between 8% and 23% over 10 screening examinations.[11-14] Specifically for the Dutch breast cancer screening programme, analyses were performed for 13 examinations which resulted in a cumulative risk of FP results ranging of 16.1%.[17] In our study, data from women starting screening in 1975 to 1976 were used and data on five screening examinations from women starting screening in 1997 were extrapolated using the data from 1975 and incorporating the expected effect of digital mammography. However, in the meantime, changes have been made in the programme such as the introduction of digital mammography, the implementation of two-view mammography in both initial and subsequent screening examinations and changes to the referral strategy including the use of the Breast Imaging-Reporting and Data System (BI-RADS) categories which affected the amount of FP results.[3,18,19]

Furthermore, international studies found that women who previously had a FP result are more likely to be diagnosed with breast cancer later on.[20-23] The reported hazard rates (HRs) and relative risks (RRs) were between 1.67 and 2.18 for women who previously had a FP result and increased to HRs between 4.22 and 9.13 for women who had multiple FP results.[20-23] Risks of both screen-detected and interval cancers were found to be increased and remained higher until 12 years after receiving the FP result.[23] This suggests that there might be some underlying biological susceptibility that causes some of the excess cancer risk in women with a FP test.[20] However, since FP rates differ between countries, it can be expected that the population of women with a history of a FP result and their risk factors are different as well. Therefore, it is unclear if, and to what extent FP results in the Dutch breast cancer screening programme lead to an increased risk of a breast cancer diagnosis. This is especially relevant since women were found to be less likely to participate in screening after a FP result in the Dutch breast cancer screening programme.[24]

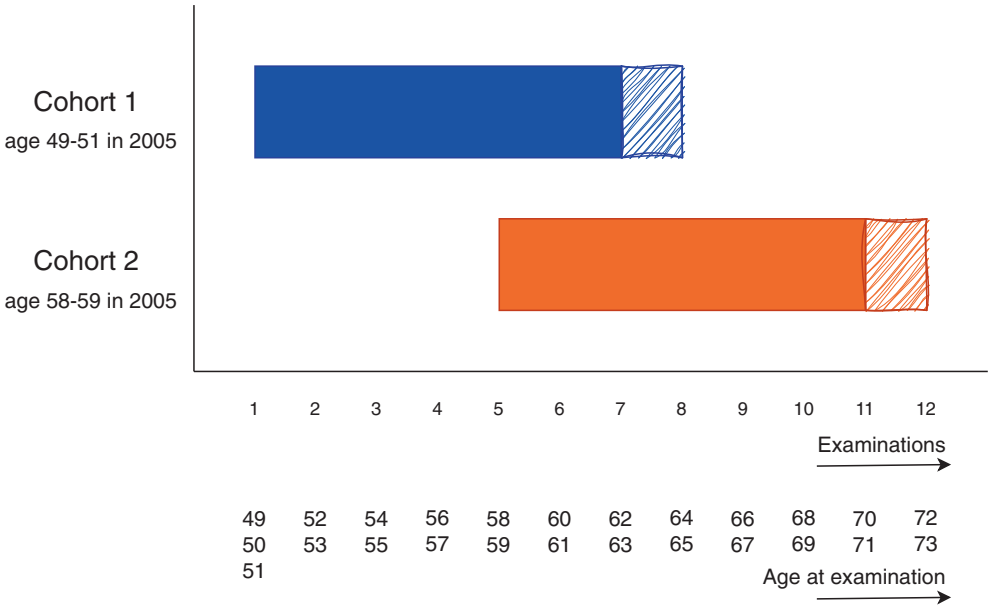

FIGURE 1 Construction of cohorts.

Therefore, our study aimed to estimate the cumulative risk of false positive recall and screen-detected breast cancer after multiple screening examinations in the Netherlands using more recent data. Furthermore, our study aimed to investigate screening behaviour and outcomes in women with a history of FP results.

2 | METHODS

The population-wide breast cancer screening programme in the Netherlands started in 1990 with biennial mammography screening for women aged 50 to 69. In 1998, this age-range was extended to also include women aged 70 to 74. Initially, screen-film mammography was used, but this was gradually replaced for full-field digital mammography between 2003 and 2010. Mammographic examinations were performed by specialised radiographers who checked the images and immediately repeat examinations in case of vagueness or incompleteness. Independent double reading is performed by specialised screening radiologists who use the BI-RADS system to classify mammograms. In the Netherlands, women with a BI-RADS score of 0, 4 and 5 are referred for follow-up testing.[18] BI-RADS 3 is not used.

2.1 | Data collection

Data were retrieved from the Dutch Cancer Registry (NKR) at the Netherlands Comprehensive Cancer Organisation (IKNL). The dataset included data on screening invitations, participation and outcomes. Furthermore, the age of the women at each screening examination was included. Participation was defined as a screening test registered after a screening invitation and before the sending of the invitation of the subsequent examination (ie, ~24 months).

At the start of the screening programme, screening data were stored in multiple regional screening registries. More recently, the data was brought together in a national database. However, due to differences in registries, data from before 2005 were incomplete which made the data unreliable for this analysis. Therefore, we chose to only include data from screening invitations sent from the year 2005 onwards.

During the time period 2005 to 2019, women who regularly received biennial breast cancer screening invitations could have been invited for breast cancer screening seven or eight times. Women who moved to another municipality in the meantime could have received more or less invitations and women who permanently unregistered for breast cancer screening or had breast cancer received less invitations.

2.2 | Population

This longitudinal, observational cohort study included two cohorts of women who were invited for breast cancer screening. The first cohort included women invited for the first time in 2005; first-time invitees. These women were either 49, 50 or 51 years of age in 2005. The second cohort included women who were 58 or 59 years of age in 2005, thus, in 2005 they were invited for breast cancer screening for the fifth or sixth time in their life. Women who did not participate after the screening invitation in 2005 were excluded from analysis. Both cohorts were followed over multiple screening examinations ending with invitations sent until December 31, 2018. Participation and result data were included until early 2020.

2.3 | Statistical analyses

Because data were only available from 2005 onwards, the analyses included seven consecutive screening examinations of the 13 that were offered in the Dutch breast cancer screening programme. However, by using the data of a second cohort of older women, extrapolation was possible until 11 examinations of screening (Figure 1).

Life table survival analyses were performed with data from the cohort of first-time invitees to evaluate the cumulative risk of receiving (1) a FP or (2) a TP screening result (ie, breast cancer diagnosis) for the first seven screening examinations. Follow-up time was censored if a woman had a TP or FP result (only first FP results were analysed), if a woman stopped participating in screening (lost-to-follow up) or when the end of the data collection was reached. Of the women who were still in the analysis after seven screening examinations, we assumed 91% of the participants to also participate for the eighth examination and therefore have a longer follow-up time.[8] Cumulative risk of having received a FP or TP result was calculated per each consecutive examination to a maximum of seven examinations of participation and presented in percentages with accompanying 95% confidence intervals (95% CI).

A similar survival analysis was performed for the second cohort of women. Assuming equal risks between the first cohort and second cohort at equal ages, the results from the survival analysis of the second cohort were used to extrapolate the cumulative risk for additional examinations. The increase in cumulative risk per examination in cohort two was applied to the cumulative risk after seven examinations in the first cohort. This allowed for extrapolation of cumulative risks up until 11 screening examinations.

Furthermore, among the first-time invitees cohort, participation in screening examinations subsequent to a true negative (TN) result was compared to participation subsequent to a FP result. In addition, the rates of breast cancer diagnoses and additional FP results were compared between women who previously received TN results and women who had a history of FP results per 1000 screens. Statistical analyses were performed in IBM SPSS Statistics version 28.

TABLE 1 Population characteristics.

	Cohort 1 First-time invitees in 2005	Cohort 2 Fifth or sixth invitation in 2005
Age range	49-51	58-59
n	92 902	66 472
Average number of invitations received between 2005 and 2019	6.9 (SD 1.1) (range 1-10)	6.8 (SD 1.1) (range 1-12)
Times participated	6.2 (SD 1.7) (range 1-9)	6.3 (SD 1.5) (range 1-9)
Result screening examination in 2005		
TP	464 (5.0 per 1000)	288 (4.3 per 1000)
FP	1998 (21.5 per 1000)	435 (6.5 per 1000)
TN	90 213 (971.1 per 1000)	65 588 (986.7 per 1000)
Interval cancer (FN)	227 (2.4 per 1000)	161 (2.4 per 1000)
FP/TP ratio	4.3	1.5

3 | RESULTS

In total, data were received from 115 122 women who received their first invitation for breast cancer screening in 2005 (cohort 1). Among them, 92 902 (80.7%) participated in this first screening examination. The second cohort consisted of 66 472 women who participated in the screening examination in 2005. After the screening examination in 2005, 97.1% of the first-time invitees and 98.7% of the older women received a TN screening result (Table 1). Furthermore, the TP rate was 5.0 and 4.3 per 1000, and the FP rate was 21.5 and 6.5 per 1000, respectively. This resulted in 4.3 FP results per TP in the first screening examinations in 2005 and 1.5 FP results per TP in the fifth or sixth examination in 2005. In both cohorts, the interval cancer (false negative [FN] screen) rate was 2.4 per 1000.

Survival analyses found that after seven screening examinations the cumulative risk of at least one FP result was 9.1% (95% CI:

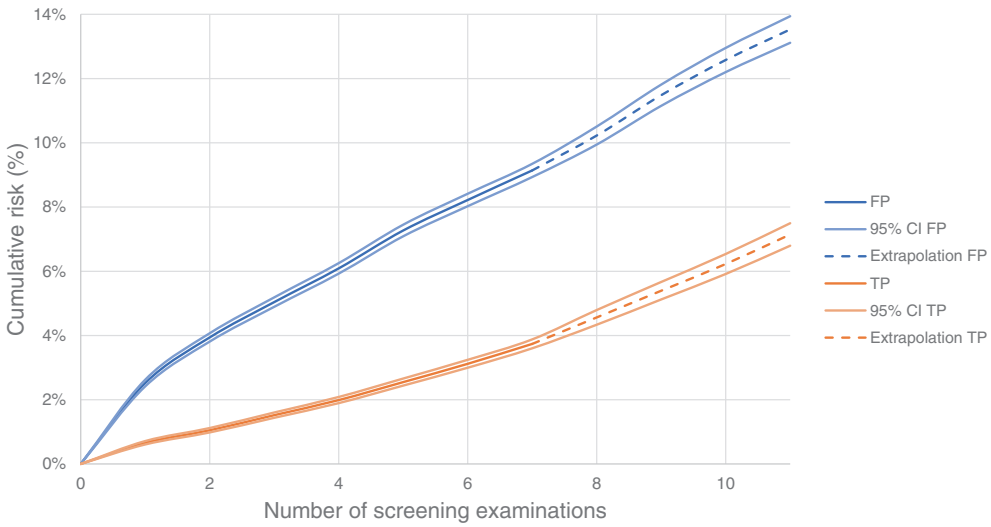

FIGURE 2 Cumulative risk of FP and TP screening results over 11 screening examinations. The dashed lines represent the extrapolation of results based on data from cohort 2.

8.9%-9.4%) and of a TP result was 3.7% (95% CI: 3.6%-3.9%) (Table S1). The extrapolated cumulative risks show an increase in cumulative risk to 13.5% (95% CI: 13.1-13.9) for FP results and 7.1% (95% CI: 6.8-7.5) for TP results after 11 screening examinations. The FP/TP ratio was highest after one or two examinations (3.8) and decreased after an increasing number of examination and an increase in age of the women (2.5 after 7 examinations and 1.9 after 11). During the first examination the highest percentage of FP results was found, after which the increase in cumulative risk seemed to follow a less steep linear trend (Figure 2). After a relative high number of TP results during the first examination, the cumulative risk increased more slowly followed by an increasing steepness during later examinations at higher age.

Participation in the screening examination following a TN screening result was found to be over 90%, independent of the examination in which the TN result was received (Figure 3). When a FP result was received in the first screening examination, participation in the second examination was 71%. However, the later the FP result was received, the higher the participation rate in the subsequent examination with a maximum of 81% participation in examinations 6 and 7. Even though the participation rate increased as the FP was received later, it was always lower than when a TN result was received.

In the screening examinations following a TN result, 34.6 per 1000 screens performed resulted in a FP result, 16.5 per 1000 in a TP result and 5.0 per 1000 in a FN result (Table 2). Resulting in a total of 21.5 breast cancer diagnosis per 1000 screens performed. Among women who previously had a FP result, per 1000 screens 67.8 were FP, 26.1 were TP and 8.2 were FN. Compared to women with TN results, women with a FP result had 1.96 times as many FP results (95% CI: 1.87-2.05), 1.59 times as many TP results (95% CI: 1.44-1.72) and 1.66 times as many FN results (95% CI: 1.41-1.91). Combining TP and FN results in total number of breast cancer diagnoses, women with a previous FP result had 1.60 times as many diagnoses per 1000 screens (95% CI: 1.48-1.72).

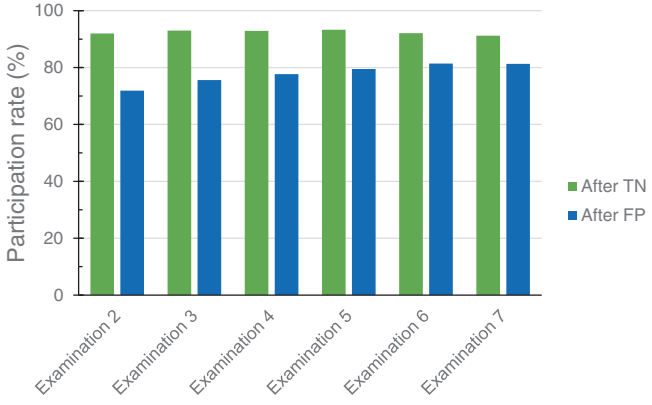

FIGURE 3 Participation in breast cancer screening stratified by previous screening result.

4 | DISCUSSION

Our study found that women who participated in the first seven screening examinations of the Dutch breast cancer screening program had a cumulative risk of 3.7% to have a screen-detected breast cancer and a 9.1% cumulative risk of having a FP screening result. This was expected to increase to 7.1% and 13.5%, respectively, after 11 screening examinations. In women who previously had a FP result, we found that the participation rate was lower than in women who had TN results. The difference between these two groups was most pronounced when women received a FP result in the first screening examination. In addition, women with a history of FP results had nearly twice as many additional FP results, and a 60% increase in breast cancer diagnoses compared to women with TN screening results. This increase in breast cancer diagnoses was observed in both screen-detected and interval cancers.

TABLE 2 FP, TP, FN and BC results in screening examinations following TN and FP results.

	After TN (516 918 screens)				After FP (7406 screens)			
	FP	TP	FN	BC (TP + FN)	FP	TP	FN	BC (TP + FN)
Examination 2	1062	270	217	487	34	4	6	10
Examination 3	1705	677	374	1051	75	15	8	23
Examination 4	2598	1100	431	1531	94	41	11	52
Examination 5	3948	1724	655	2379	85	48	16	64
Examination 6	4037	2284	633	2917	100	40	13	53
Examination 7	4521	2474	254	2728	114	45	7	52
Total	17 871	8529	2564	11 093	502	193	61	254
/1000	34.6	16.5	5	21.5	67.8	26.1	8.2	34.3
95% CI	*34.1-35.1*	*16.2-16.8*	*4.8-5.2*	*21.1-21.9*	*62.1-73.5*	*22.4-29.7*	*6.2-10.3*	*30.2-38.4*
Rate FP/rate TN					1.96	1.59	1.66	1.6
95% CI					*1.87-2.05*	*1.44-1.72*	*1.41-1.91*	*1.48-1.72*

Abbreviation: 95% CI, 95% confidence interval.

A study in Finland found a cumulative risk of a screen-detected breast cancer of 3.4% over 7 screening examinations and 5.7% over 10 examinations with the highest risk in women with a history of breast cancer symptoms.[11] Furthermore, a Spanish study found that women with a history of benign breast disease had a cumulative risk of 3.6%, women with a family history of breast cancer had a cumulative risk of 4.5%, women with both had a risk of 6.1% and women with neither had a cumulative risk of 2.6% over seven screening examinations.[25] The weighted average of these four groups would come down to a cumulative risk of 3.0%. Compared to our results on TPs, the Finnish and Spanish risks are a little lower. A reason for this is the lower breast cancer incidence in both countries compared to the Netherlands.[26] However, also differences in screening detection performance may play a role.[27] A previous study on the Dutch breast cancer screening programme predicted that the cumulative risk of a screen-detected breast cancer after implementation of digital mammography would be 7.1% over 13 examinations of screening.[17] Our study already found a cumulative risk of 7.1% after 11 examinations. The increase in cumulative risk can probably be explained by the usage of data from a more recent cohort of women who have a higher risk of developing breast cancer.[28]

Only a few studies present the cumulative risk of a FP result after seven examinations of breast cancer screening. A study in Spain found cumulative risks between 20.7% and 34.3% depending on family history and previous benign breast disease, an Italian study found a cumulative risk of 15.2%, and a Finnish study found a cumulative risk of 13.6% to receive a FP result after seven examinations.[11,25,29] All three estimates are higher than the cumulative risk of 9.1% that we found after seven examinations in the current study. More European studies reported cumulative risks after 10 screening examinations and found estimates between 8% and 23%.[11-14] Only the cumulative risk of 8% found in the region of Fyn in the Danish study was lower than the 12.6% the current study found after 10 examinations of screening.[14] The estimate for the Copenhagen region and the other studies were all higher than the 12.6% we found over 10 examinations and also higher than the 13.5% we found for 11 screening examinations. In addition, two American studies found even higher cumulative risks between 38.1% and 42% after five examinations of biennial screening and between 56.3% and 61.3% after 10 examinations of annual screening.[15,16] The considerable difference between most European countries and the US can probably be explained because of the lower recall rate in most European countries compared to the US.[7] Even within Europe, recall rates differ and can explain the differences between countries, but also differences in calendar year of data used could have an influence.[27] Additionally, the study by Ho et al found that screening with digital breast tomosynthesis (DBT) instead of digital mammography can decrease the cumulative risk of a FP results by 6.7% point in annual and 2.4% point in biennial screening.[16] In the US DBT is used in a proportion of the screening settings, while in Europe DBT is hardly used in screening, which can also explain part of the difference between the estimates in the US and Europe.

Overall, the cumulative risks on FP and TP results in the Netherlands are relatively favourable compared to other countries. Despite the lower recall rate in the Netherlands, the cumulative risk of a screen-detected breast cancer remains quite comparable. This is an indication that the lower recall rate did not compromise the detection rate. This was also reflected in the low FP/TP ratios found, compared to a FP/TP ratio of 3.2 after 10 examinations in Finland and of the pooled estimate of 2.8 presented in the "balance sheet" by Paci et al based on European data and studies.[11,30] Especially considering the usage of digital mammography in the majority of screens performed in the current study which was usually found to yield a higher FP rate.[31] In addition, it shows that the Dutch policy of not including follow-up diagnostic assessment as part of the screening programme did not compromise the cumulative FP and TP rates. This low FP/TP ratio is expected to translate into a favourable ratio in long-term harms and benefits.

Previously reported HRs and RRs for breast cancer diagnosis after a FP result were between 1.7 and 2.2 for women who previously had a FP result and increased to 4.2 to 9.1 for women who had multiple FP results.[20-23] The increased incidence ratio in our study was 1.59 for screen-detected cancers and 1.66 for interval cancers which is in line with the lower rates found in the previous studies. However, some studies found that part of the FP results were misclassified because of a false negative diagnostic assessment (FNDA) which would be the case for 0.6% to 1.5% of recalled women.[20,21] After exclusion of these women, the HR in Flanders decreased from 1.9 to 1.5 and the RR in Denmark decreased from 1.7 to 1.3.[20,21] It is unclear whether FNDA occurs in the Netherlands, and if so, to what extent this happens. Unfortunately, the data required to investigate this was unavailable. Therefore, it was not possible to adjust for this in the current analysis.

Participation among women with a TN result was found to be high, over 90%, which is in line with the participation loyalty in the monitors of the Dutch screening programme.[8] Among women with a TP result, participation was found to be lower. This was also found in two previous studies in the Netherlands which found even lower participation rates of around 65% among women with a FP result compared to 93% to 95% among women with a negative screening result.[9,24] However, Setz-Pels et al also found that nearly 30% of women with a FP result had follow-up surveillance in the hospital, which suggested that the mammography coverage, that is, screening coverage and hospital surveillance combined, in women with a history of FP results would be almost as high as for women with a negative screening result.[9] On the other hand, a study in Copenhagen did not find any difference between women with a negative and women with a FP screening result in their participation rate in the next round.[32] Interestingly, Chiarelli et al found that, in Ontario, re-attendance of previously FP women was lower in screening centres without an assessment programme, like the policy in the Netherlands, and equal to negative women in centres with an assessment programme, like the policy in Denmark.[33] Since screening centres in the Netherlands do not have assessment included, while in Denmark assessment is part of the screening programme, this might explain the difference in re-attendance behaviour.

A strength of our study was the use of registry data, which included practically all screening invitations and tests performed in

the Netherlands. Unfortunately, data from before 2005 were incomplete, which restricted the data to only include seven screening examinations of the 13 that were offered in the Dutch breast cancer screening programme. However, by using the data of the second cohort of older women, extrapolation was possible until 11 examinations of screening. This extrapolation was performed under the assumption that both cohorts were equal in breast cancer risk at the screening examination extrapolated. This assumption largely holds, because during those screening examinations, the women in both cohorts would have had the same age. This was confirmed by comparable increases in cumulative risk of FP and TP results between examinations 5 and 7, of which data were available for both cohorts. A difference in risk could have been caused by a difference in birth cohort risk as shown by van der Waal et al and Napolitano et al, but this effect was expected to be relatively small since the age difference was only 9 years.[28,34] Even though the extrapolation is less precise than analysis based on observed data, the benefit of this was that the cumulative risks are more applicable to women eligible for screening in current times, because screening performance has changed due to the implementation of digital mammography and because the breast cancer risk has increased over the years.[28,35]

Given that most women in the Netherlands seem to make a fundamental decision about participation in breast cancer screening and adhere to this decision for future invitations, it is important that this decision is based on information encompassing benefits and harms of participation in multiple screening examinations. In addition, providing stratified information on increased risks in women who previously had a FP outcome may give them insights into their personal risk of developing breast cancer and may potentially increase their participation to the screening programme. Furthermore, in the prospect of risk stratified screening, it may be useful to include history of FP results into consideration when forming risk groups.

To conclude, we found that women who participate in the Dutch breast cancer screening programme have a cumulative risk of 3.7% to receive a TP result and of 9.1% to receive a FP result after seven screening examinations. After 11 examinations, these risks would increase to 7.1% for a TP result and 13.5% for a FP result. Due to the low recall rate in the Netherlands, these cumulative risks are relatively favourable compared to screening programmes in other countries, which is also represented in a favourable FP/TP ratio which is expected to translate in a favourable ratio in long-term harms and benefits. Furthermore, women who previously received a FP result more often receive TP and FP results in later examinations and more often have interval cancers while their participation rate in subsequent examinations is lower compared to women with TN results.

AUTHOR CONTRIBUTIONS

Lindy M. Kregting: Conceptualisation, methodology, formal analysis, investigation, writing - original draft, visualisation; **Nicolien T. van Ravesteyn:** Conceptualisation, methodology, validation, writing - review & editing, supervision, funding acquisition; **Sarocha Chootipongchaivat:** Conceptualisation, Methodology, validation, writing - review & editing; **Eveline A. M. Heijnsdijk:** Conceptualisation, validation, writing - review & editing; **Johannes D. M. Otten:** Conceptualisation, validation, writing - review & editing; **Mireille J. M. Broeders:** Conceptualisation, validation, writing - review & editing, funding acquisition; **Harry J. de Koning:** Conceptualisation, writing - review & editing, supervision, funding acquisition. The work reported in the article has been performed by the authors, unless clearly specified in the text.

ACKNOWLEDGEMENTS

We would like to thank Dr. Daan Nieboer for discussions on the statistical analysis methods.

FUNDING INFORMATION

This research was funded by the Centre for Population Screening (CvB) of the Dutch National Institute for Public Health and the Environment (RIVM). The CvB was not in any way involved in the study design, analysis or interpretation of the data, nor in the writing of the report or in the decision to submit the article for publication.

CONFLICT OF INTEREST STATEMENT

Dr. van Ravesteyn reports receiving fees for consulting from Wickenstones (paid to institution). Prof. dr. de Koning declares receiving fees for lectures at symposia for TEVA, Menarini, and Astra Zeneca (<€1000) and for reviewing external model analyses for Bayer (<€1000). The other authors have no conflicts to declare.

DATA AVAILABILITY STATEMENT

The data used in our study are from the Netherlands Cancer Registry (NCR) and property of the Netherlands Comprehensive Cancer Organisation (IKNL). Data from the NCR can be requested at IKNL via https://iknl.nl/en/ncr/apply-for-data. Further information is available from the corresponding author upon request.

ETHICS STATEMENT

Ethical approval was not applicable. In the Netherlands, implicit informed consent to pseudonymised cancer screening data use in research is recorded when one takes up the offer of screening. All individuals were informed that they could explicitly withdraw consent.

ORCID

Lindy M. Kregting https://orcid.org/0000-0002-8262-0080
Sarocha Chootipongchaivat https://orcid.org/0000-0002-7714-4737
Eveline A. M. Heijnsdijk https://orcid.org/0000-0002-4890-6069
Mireille J. M. Broeders https://orcid.org/0000-0002-8741-8148
Harry J. de Koning https://orcid.org/0000-0003-4682-3646

REFERENCES

1. Zielonke N, Gini A, Jansen EEL, et al. Evidence for reducing cancer-specific mortality due to screening for breast cancer in Europe: a systematic review. *Eur J Cancer*. 2020;127:191-206.
2. Ponti A, Anttila A, Ronco G, Senore C. *Cancer Screening in the European Union*. International Agency for Research on Cancer; 2017.

3. Sankatsing VDV, Geuzinge HA, Fracheboud J, et al. Landelijke evaluatie van bevolkingsonderzoek naar borstkanker in Nederland 2004-2014. 2019.
4. Ripping TA-O, Verbeek AL, Fracheboud J, de Koning HJ, van Ravesteyn NT, Broeders MJ. Overdiagnosis by mammographic screening for breast cancer studied in birth cohorts in The Netherlands. *Int J Cancer*. 2015;137:921-929.
5. Bond M, Pavey T, Welch K, et al. Psychological consequences of false-positive screening mammograms in the UK. *Evid Based Med*. 2013;18:54-61.
6. Williams J, Garvican L, Tosteson AN, Goodman DC, Onega T. Breast cancer screening in England and the United States: a comparison of provision and utilisation. *Int J Public Health*. 2015;60:881-890.
7. Kemp Jacobsen K, O'Meara ES, Key D, et al. Comparing sensitivity and specificity of screening mammography in the United States and Denmark. *Int J Cancer*. 2015;137:2198-2207.
8. Monitor bevolkingsonderzoek borstkanker 2019. 2021 https://www.rivm.nl/documenten/monitor-bevolkingsonderzoek-borstkanker-2019
9. Setz-Pels W, Duijm LE, Coebergh JW, Rutten M, Nederend J, Voogd AC. Re-attendance after false-positive screening mammography: a population-based study in The Netherlands. *Br J Cancer*. 2013;109:2044-2050.
10. 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*. 2011;22:925-929.
11. Singh D, Pitkaniemi J, Malila N, Anttila A. Cumulative risk of false positive test in relation to breast symptoms in mammography screening: a historical prospective cohort study. *Breast Cancer Res Treat*. 2016;159:305-313.
12. Roman M, Skaane P, Hofvind S. The cumulative risk of false-positive screening results across screening centres in the Norwegian Breast Cancer Screening Program. *Eur J Radiol*. 2014;83:1639-1644.
13. Salas D, Ibáñez J, Román R, et al. Effect of start age of breast cancer screening mammography on the risk of false-positive results. *Prev Med*. 2011;53:76-81.
14. Njor SH, Olsen AH, Schwartz W, Vejborg I, Lynge E. Predicting the risk of a false-positive test for women following a mammography screening programme. *J Med Screen*. 2007;14:94-97.
15. Hubbard RA, Miglioretti DL. A semiparametric censoring bias model for estimating the cumulative risk of a false-positive screening test under dependent censoring. *Biometrics*. 2013;69:245-253.
16. Ho T-QH, Bissell MCS, Kerlikowske K, et al. Cumulative probability of false-positive results after 10 years of screening with digital breast Tomosynthesis vs digital mammography. *JAMA Netw Open*. 2022;5:e222440.
17. Otten JDM, Fracheboud J, den Heeten GJ, et al. Likelihood of early detection of breast cancer in relation to false-positive risk in life-time mammographic screening: population-based cohort study. *Ann Oncol*. 2013;24:2501-2506.
18. Timmers JMH, van Doorne-Nagtegaal HJ, Zonderland HM, et al. The breast imaging reporting and data system (BI-RADS) in the Dutch breast cancer screening programme: its role as an assessment and stratification tool. *Eur Radiol*. 2012;22:1717-1723.
19. Blanks RG, Bennett RL, Patnick J, Cush S, Davison C, Moss SM. The effect of changing from one to two views at incident (subsequent) screens in the NHS breast screening programme in England: impact on cancer detection and recall rates. *Clin Radiol*. 2005;60:674-680.
20. von Euler-Chelpin M, Kuchiki M, Vejborg I. Increased risk of breast cancer in women with false-positive test: the role of misclassification. *Cancer Epidemiol*. 2014;38:619-622.
21. Goossens MC, de Brabander I, de Greve J, et al. Breast cancer risk is increased in the years following false-positive breast cancer screening. *Eur J Cancer Prev*. 2017;26:396-403.
22. Castells X, Tora-Rocamora I, Posso M, et al. Risk of breast cancer in women with false-positive results according to mammographic features. *Radiology*. 2016;280:379-386.
23. Roman M, Hofvind S, von Euler-Chelpin M, Castells X. Long-term risk of screen-detected and interval breast cancer after false-positive results at mammography screening: joint analysis of three national cohorts. *Br J Cancer*. 2019;120:269-275.
24. Klompenhouwer EG, Duijm LEM, Voogd AC, et al. Re-attendance at biennial screening mammography following a repeated false positive recall. *Breast Cancer Res Treat*. 2014;145:429-437.
25. Román M, Quintana MJ, Ferrer J, Sala M, Castells X. Cumulative risk of breast cancer screening outcomes according to the presence of previous benign breast disease and family history of breast cancer: supporting personalised screening. *Br J Cancer*. 2017;116:1480-1485.
26. Dafni U, Tsourti Z, Alatsathianos I. Breast cancer statistics in the European Union: incidence and survival across European countries. *Breast Care*. 2019;14:344-353.
27. Armaroli P, Riggi E, Basu P, et al. Performance indicators in breast cancer screening in the European Union: a comparison across countries of screen positivity and detection rates. *Int J Cancer*. 2020;147:1855-1863.
28. van der Waal D, Verbeek ALM, den Heeten GJ, Ripping TM, Tjan-Heijnen VCG, Broeders MJM. Breast cancer diagnosis and death in The Netherlands: a changing burden. *Eur J Public Health*. 2015;25:320-324.
29. Puliti D, Miccinesi G, Zappa M. More on screening mammography. *N Engl J Med*. 2011;364:281-286.
30. Paci E. Summary of the evidence of breast cancer service screening outcomes in Europe and first estimate of the benefit and harm balance sheet. *J Med Screen*. 2012;19:5-13.
31. Farber R, Houssami N, Wortley S, et al. Impact of full-field digital mammography versus film-screen mammography in population screening: a meta-analysis. *J Natl Cancer Inst*. 2021;113:16-26.
32. Andersen SB, Vejborg I, von Euler-Chelpin M. Participation behaviour following a false positive test in the Copenhagen mammography screening programme. *Acta Oncol*. 2008;47:550-555.
33. Chiarelli AM, Mai V, Moravan V, Halapy E, Majpruz V, Tatla RK. False-positive result and reattendance in the Ontario breast screening program. *J Med Screen*. 2003;10:129-133.
34. Napolitano G, Lynge E, Lillholm M, et al. Change in mammographic density across birth cohorts of Dutch breast cancer screening participants. *Int J Cancer*. 2019;145:2954-2962.
35. de Munck L, de Bock GH, Otter R, et al. Digital vs screen-film mammography in population-based breast cancer screening: performance indicators and tumour characteristics of screen-detected and interval cancers. *Br J Cancer*. 2016;115:517-524.

SUPPORTING INFORMATION

Additional supporting information can be found online in the Supporting Information section at the end of this article.

How to cite this article: Kregting LM, van Ravesteyn NT, Chootipongchaivat S, et al. Cumulative risks of false positive recall and screen-detected breast cancer after multiple screening examinations. *Int J Cancer*. 2023;153(2):312-319. doi:10.1002/ijc.34530

B-cell malignancies - A new knowledge hub on the latest research in therapeutic advances

EDUCATIONAL CONTENT AVAILABLE ON THE HUB:

- On-demand Webinars - earn CME credit
- Infographics
- Patient Case Studies
- Currated Research Articles

...and much more

VISIT KNOWLEDGE HUB TODAY

This educational resource has been supported by Eli Lilly.

WILEY