

# Impact of mammographic screening on the detection of good and poor prognosis breast cancers

Laura J. Esserman · Yiwey Shieh · Emiel J. T. Rutgers · Michael Knauer ·  
Valesca P. Retèl · Stella Mook · Annuska M. Glas · Dan H. Moore ·  
Sabine Linn · Flora E. van Leeuwen · Laura J. van 't Veer

Received: 15 August 2011 / Accepted: 17 August 2011 / Published online: 4 September 2011  
© Springer Science+Business Media, LLC. 2011

**Abstract** We sought to compare the molecular signature of node negative cancers from two cohorts 15 years apart, to determine if there is molecular evidence of increase in low and ultralow risk cancers over time. We studied the impact of age, time period of diagnosis, and mammographic screening on biology of tumors where The Netherlands Cancer Institute 70-gene prognosis signature was generated as part of 2 validation series, one retrospective (1984–1992), Cohort 1, and one prospective (2004–2006), Cohort 2. A total of 866 patients were analyzed. Regardless of time period of diagnosis, the proportion of T1, grade 1, hormone receptor positive (HR) tumors, and good prognosis by 70-gene signature significantly increases as age increases ( $P < 0.01$ ). In women aged 49–60, the time period of diagnosis significantly affects the proportion of cancers that

were NKI 70-gene low risk: 40.6% (67/165) compared with 58% (119/205) for Cohorts 1 and 2, respectively. This is in contrast to the absence of a significant change for women under age 40, where 25% (17/68) and 30% (17/56) were low risk in Cohorts 1 and 2, respectively. In women aged 49–60, using an ultralow risk threshold of the 70-gene signature, 10% of tumors in Cohort 1 were ultralow risk compared with 30% for women with screen-detected cancers in Cohort 2. Older age and method of detection (screening) are associated with a higher likelihood of a biologically low risk tumor. In women over age 50, biologically low risk tumors are frequent and tools that classify risk may minimize overtreatment.

**Keywords** Breast cancer · Mammographic screening · NKI 70-gene prognosis signature · Node negative cancers

L. J. Esserman (✉) · Y. Shieh · D. H. Moore  
University California San Francisco (UCSF),  
San Francisco, CA, USA  
e-mail: Laura.esserman@ucsfmedctr.org

Y. Shieh  
e-mail: Yiwey.Shieh@ucsf.edu

E. J. T. Rutgers · M. Knauer · V. P. Retèl · S. Mook · S. Linn ·  
F. E. van Leeuwen · L. J. van 't Veer  
The Netherlands Cancer Institute, Amsterdam, The Netherlands  
e-mail: f.v.leeuwen@nki.nl

L. J. van 't Veer  
e-mail: vantveerl@cc.ucsf.edu

A. M. Glas  
Agendia NV, Amsterdam, The Netherlands

*Present Address:*

L. J. van 't Veer  
University California San Francisco (UCSF),  
San Francisco, CA, USA

## Introduction

There has been a sustained increase in the incidence of breast cancer largely due to node negative cases in women over 50 [1] which has been ascribed to the introduction of population-based mammographic screening programs as well as other shifts in population risk factors. A concomitant decrease in regional or node positive disease has not been observed. This phenomenon has been seen in many countries that have instituted population-based screening programs. Whether this reflects an increase in the incidence of low risk tumors or the detection of clinically indolent disease (over diagnosis) [2], or a combination of both, is not clear.

Early detection makes a difference in outcome for some cancers. However, critical questions remain about the impact of mammography on the increased detection of low risk tumors. If indeed much of the increase in incidence is

second to tumors with low malignant potential and if such cancers can be reliably identified, then patients and providers could make informed decisions to avoid treatments that may do more harm than good.

A number of tools for predicting the metastatic potential of breast cancers are now available. These range from tools that predict risk based on standard clinicopathologic features such as ADJUVANT! [3] (adjuvantonline.com) to molecular tools that are available for research [4] and commercial use in breast cancer [5, 6]. These tools provide the opportunity to study the biologic presentations of breast cancers before and after the introduction of population-based mammographic screening. We took advantage of datasets collected before and after the introduction of population-based screening where the 70-gene prognosis signature (now marketed as MammaPrint™) results were available to further investigate whether the introduction of screening has influenced the proportion of biologic characteristics of node negative detected breast cancers.

The NKI 70-gene prognosis signature was developed to predict long-term outcome in the absence of systemic therapy on a consecutive series of breast cancer patients from The Netherlands Cancer Institute (NKI). Using a classification of good versus poor prognosis (low vs. high risk), it was found to be predictive of overall survival and development of distant metastases [7]. The NKI 70-gene prognosis signature was validated in a cohort of patients from five European centers in the TRANSBIG study, where over a median follow-up period of 13.6 years, a poor prognosis signature was associated with a hazard ratio (HR) for time to distant metastasis of 2.32 (95% CI = 1.35–4.00), and a HR of 2.79 (95% CI = 1.6–4.87) for overall survival [8]. The Dutch Health Insurance Board, in 2003, sponsored a prospective study, known as Microarray Prognostics in Breast Cancer (RASTER), to determine the feasibility of integrating the 70-gene prognosis signature into routine care at community and academic settings in the Netherlands [9].

Prior to 1992, in Europe, population-based mammography was not routinely offered, and screening uptake rates were 25% or less [3]. Thus, the patients diagnosed before 1992 from the NKI and TRANSBIG data reflect the biology of node negative tumors detected before the use of population-based screening. The RASTER dataset, however, allows us to compare the distribution of the 70-gene index in node negative tumors detected after population-based screening was introduced with compliance reaching 80% [10].

In the current study, we examine the 70-gene signatures in node negative cancers in these two cohorts which differ by an era of diagnosis, as well as screening uptake rates. In the cohort representing the earlier era, screening was not routine, while screening was routine in the later cohort. These data sets allow us to answer two questions. First, has

**Table 1** Patient cohorts

	Cohort 1	Cohort 2
Time period	1984–1992	2004–2006
Age at diagnosis	All ages	Up to age 60 <sup>a</sup>
Screening routine	No	Yes (49–60)
Recruitment	NKI and TRANSBIG	Netherlands (16 hospitals)

Composition of patients diagnosed with lymph-node negative breast cancer ([7, 8, 12] vs. [9] respectively) in Cohorts 1 and 2

<sup>a</sup> Enrollment RASTER up to age 60

there been a shift toward the detection of molecularly good prognosis cancers? Second, is there an increase in the detection of ultralow risk tumors, those that may have an excellent prognosis in the absence of systemic treatment, and are these cancers more common in the screen-detected group? Since age is a well-established influence on the biology of cancers, these comparisons were stratified by age. Our rich data sets with molecular tumor profiles provide a unique opportunity to answer these questions.

## Methods

### Patients

Patients were selected from a database at the NKI containing patient and clinicopathologic characteristics, as well as 70-gene signature results, for 1,696 participants in previously reported studies on the 70-gene signature [11]. Two cohorts were analyzed, as summarized in Table 1.

Cohort 1 is comprised of patients whose cancers were diagnosed in an era before the widespread implementation of mammographic screening. A total of 439 node negative patients were selected from previous studies by the NKI [6, 12] and TRANSBIG consortium [8]. Patients from the TRANSBIG consortium [8] were treated at centers in England, France, and Sweden, between 1980 and 1998, and at NKI between 1984 and 1998. We restricted our analysis here to patients diagnosed before 1992. During this time period, population-wide screening was not offered routinely in these countries. We estimate that the screening uptake in the cases in this cohort was 25% or lower [13]. Cohort 1 included 68 patients diagnosed under age 40 (25 NKI and 43 TRANSBIG), 141 ages 40–48 (45 NKI and 96 TRANSBIG), 165 ages 49–60 (51 NKI and 114 TRANSBIG), and 65 who were over age 60 (65 NKI).

Cohort 2 is comprised of patients whose cancers were diagnosed in an era of widespread mammographic screening with modern equipment and techniques and at a time when women were much more likely to perform routine self-breast exam. A total of 427 patients with node

negative cancers were selected from a community-based feasibility study of the 70-gene prognosis signature for patients up to age 60 (RASTER). Of these, 56 were under 40, 166 were 40–48, and 205 patients were aged 49–60. Patients were diagnosed and treated from 2004 to 2006 at 16 centers in the Netherlands during which screening uptake approached 80% [13].

Information about method of detection (i.e., screen-detected vs. interval or non-screenings-related carcinoma) was retrieved from the medical registries at the participating hospitals and originated from the national screening facilities. Clinicopathological characteristics were available from the original publications.

### 70-Gene prognosis signature

The 70-gene prognosis signature was originally presented both as a binary result, good versus poor prognosis, and as an index score, from  $-1$  to  $1$  [6, 7]. Patients with an index score greater than  $0.4$  are classified as having a good prognosis (low risk), and those with a score less than  $0.4$  are classified as having a poor prognosis (high risk) [7]. This threshold defines the commercial test. A second threshold, an index score of  $0.6$ , was identified where no distant metastases were observed at 5 years in the group of

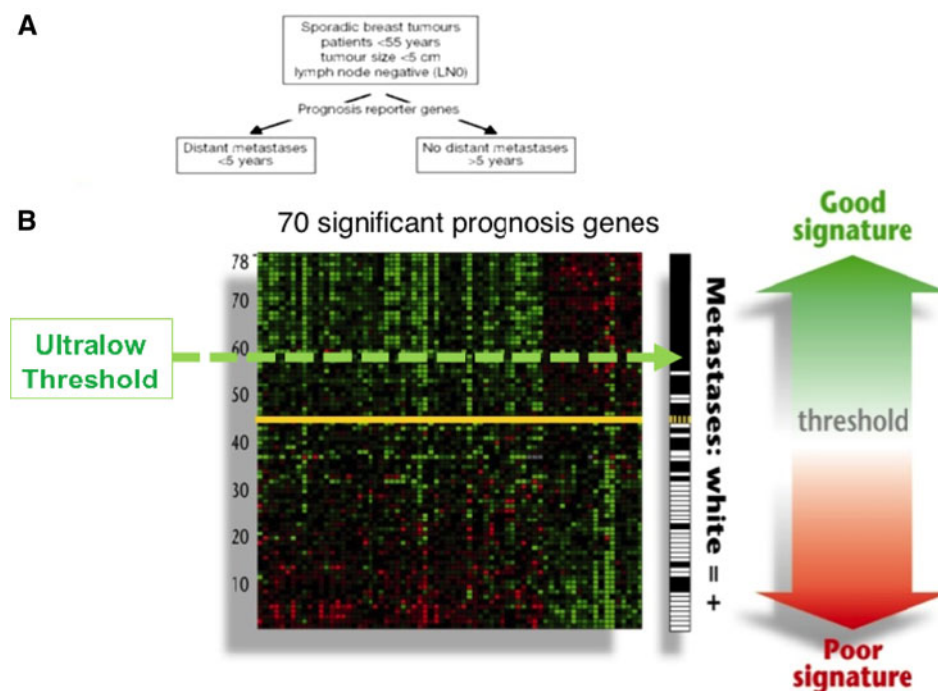
the original 78 patients [7]. The index results above  $0.6$  are referred to as the ‘ultralow’ risk range (Fig. 1).

### Statistical analysis

For Cohorts 1 and 2 combined, as well as separate, age stratified analyses were performed for women aged under 40, 40–48, 49–60 years, and over 60 years.

Median age of patients was compared using the Mann–Whitney–Wilcoxon (MWW) test, while the distributions of tumor stage and grade were compared using the Cochran–Armitage linear trend in proportions test, tumor stage 1 versus 2–4 combined and grade 1 versus 2–3 combined, respectively. The Pearson  $\chi^2$  test was used to compare the percentage of estrogen-receptor positive tumors, as well as the proportion of molecular low and high-risk tumors, between patients diagnosed in Cohorts 1 and 2. Reported  $P$  values are two-sided.

To adjust for possible differences in median age between Cohorts 1 and 2 within age strata, a multivariate logistic regression model adjusting for age was constructed for comparison of 70-gene signature between Cohorts 1 and 2. As the 70-gene signature has been previously shown to be independently associated with grade and hormone receptor (HR) status, these variables were not included in the model [8].



**Fig. 1** Identification of an ultralow risk subset. An additional 70-gene signature index score was designated as ultra-low (threshold at index score =  $0.6$ ). Expression array heat map showing the 70-gene profile for the original 78 patients. Every row represents a patient and every column one of the 70 genes. The standard threshold for good prognosis tumors is represented by the *thick red dashed line* and the threshold for

the ultralow risk designation is the *thin blue dashed line*. Adjacent to the array on the right is the Cosine correlation coefficient to the average good prognosis profile and represents the index score. The column on the far right shows the outcome for each patient either *black* (absence of distant metastasis) or *white* (presence of metastasis). Adapted from van't Veer, Nature [7]

For women diagnosed between ages 49–60 years, the value distributions of the index scores were compared among Cohort 1, Cohort 2, and patients from Cohort 2 with screen-detected cancers only (Cohort 2<sub>SD</sub>). Significance was tested using a MWW test with index score as a continuous variable. MWW analysis was also used to test for a difference in age distribution between these groups.

Analyses were performed using PASW Statistics 18.0 (SPSS Inc., Chicago, IL).

## Results

Cohort 1 and 2 represent patients diagnosed before routine mammographic screening was introduced, and when

**Table 2** Characteristics of tumor type by age across both patient cohorts

Age:	<40		40–48		49–60		>60		
Total <i>N</i>	124		307		370		65		866
T stage									
T1	64	52%	189	62%	242	65%	38	59%	<i>P</i> = 0.01*
T2	59	47%	118	38%	127	35%	26	40%	
T3	1	1%	0	0%	1	0%		0%	
Grade									
I	13	10%	58	19%	91	25%	21	32%	<i>P</i> < 0.01
II	41	33%	111	36%	161	44%	27	42%	
III	69	56%	137	45%	114	31%	17	26%	
HR-positive	74	60%	232	76%	297	80%	52	80%	<i>P</i> < 0.01
HR-negative	50	40%	75	24%	73	20%	13	20%	
NKI 70-gene									
Good prognosis	34	27%	138	45%	186	50%	41	63%	<i>P</i> < 0.01
Poor prognosis	90	73%	169	55%	180	50%	24	37%	

\* *P* value for T1 versus T2–3

**Table 3** Characteristics of tumor type by patient cohort and by age

Age:	<40		40–48		49–60		>60							
	Cohort 1	Cohort 2	Cohort 1	Cohort 2	Cohort 1	Cohort 2	Cohort 1	Cohort 2	Cohort 1	Cohort 2				
Total ( <i>n</i> )	68		141		166		165		205		65			
Median age at diagnosis (years)	35.4		44		45.3		54		52.2		65.4			
T stage	<i>n</i>	(%)	<i>n</i>	(%)	<i>n</i>	(%)	<i>n</i>	(%)	<i>n</i>	(%)	<i>n</i>	(%)		
T1	30	45	34	61	75	53	114	69	89	54	153	75	38	59
T2	38	55	21	38	66	47	52	31	75	45	52	25	26	41
T3	0	0	1	2	0	0	0	0	1	1	0	0	0	0
Comparison	<i>P</i> = 0.015				<i>P</i> = 0.006				<i>P</i> = 0.001					
Grade														
I	5	7	8	14	24	17	34	20	29	18	62	30	21	32
II	21	32	20	36	48	34	63	38	71	43	90	44	27	42
III	41	59	28	50	68	49	69	42	61	37	53	26	17	26
Comparison	<i>P</i> = 0.321				<i>P</i> = 0.775				<i>P</i> = 0.009					
Hormone receptor														
Positive	38	57	36	64	100	71	132	80	123	75	174	85	52	80
Negative	30	43	20	36	41	29	34	20	42	25	31	15	13	20
Comparison	<i>P</i> = 0.282				<i>P</i> = 0.109				<i>P</i> = 0.012					
70-gene														
Good prognosis	17	25	17	30	55	39	83	50	67	41	119	58	41	63
Poor prognosis	51	75	39	70	86	61	83	50	98	59	86	42	24	37
Comparison	<i>P</i> = 0.506				<i>P</i> = 0.054				<i>P</i> = 0.012					

Clinicopathologic characteristics of study patients by age group are shown for each cohort

compliance had reached 80%, respectively (Table 1, “Methods” section). The clinicopathologic characteristics of all patients included in this study are shown by age in Table 2. As women age, there is a significant shift toward lower T stage, grade, HR positivity status, and 70-gene good prognosis. 73% of tumors in women under 40 are 70-gene poor prognosis, compared with 50% for women aged 49–60 and 37% for women diagnosed over age 60.

Table 3 shows the characteristics of the patients by cohort. For young women, under the age of 40, tumors were found at an earlier stage (lower T stage) in the second time period compared with those diagnosed in the first cohort (1980–1991), although there was no difference in the distribution of grade, HR+ fraction, or percentage of 70-gene good prognosis tumors. Age 49 was chosen as the lower bound for the category of women aged 49–60 as 49 years is the age when women are first invited for screening in the Netherlands. Remarkably, for patients aged 49–60, Cohort 2 (2003–2006) patients had significantly higher percentages of T1 tumors, more favorable tumor grade distribution, a higher percentage of HR positive, and more 70-gene good prognosis tumors. For the 40–48 year age group, there is also a shift to a smaller proportion of poor prognosis tumors in Cohort 2, 61.3–50.0% ( $P = 0.054$ ) for Cohorts 1 and 2, respectively.

Figure 2 shows the relative percentages of good and poor prognosis tumors as defined by the established threshold for the 70-gene prognosis signature for each of the Cohorts for the age group <40 and 49–60 years. In the under 40 age group (Fig. 2a), the percentage of poor prognosis cancers was significantly higher for both cohorts. However, the distribution of poor versus good prognosis cancers remained similar regardless of period of diagnosis ( $P = 0.506$ ): 75.0 and 70.0% of cancers in the pre-screening and modern screening eras, respectively, had a poor prognosis.

In Cohort 1, for patients age 49–60, 59.4% were poor prognosis signature compared with 42.0% in Cohort 2,  $P = 0.012$  (Fig. 2b). In Cohort 2, the data is presented for the overall cohort as well as for women who presented with screen-detected cancers (Cohort 2<sub>SD</sub>), for whom the percentage of poor prognosis cancers decreased further to 33.0% ( $P < 0.01$ , compared with Cohort 1). There was a statistically significant difference in median age between Cohorts 1 and 2, with the recent cohort younger, even within the 49–60 year age group, which could have diluted the effect. The difference in 70-gene signature risk score remained significant after adjusting for the effect of age in a multivariate logistic regression model. Note that 49% of non-screen-detected cancers in Cohort 2 were poor prognosis by 70-gene profile.

To determine whether the higher fraction of good prognosis tumors was due to enrichment for cancers with the most favorable prognosis, the distribution of 70-gene

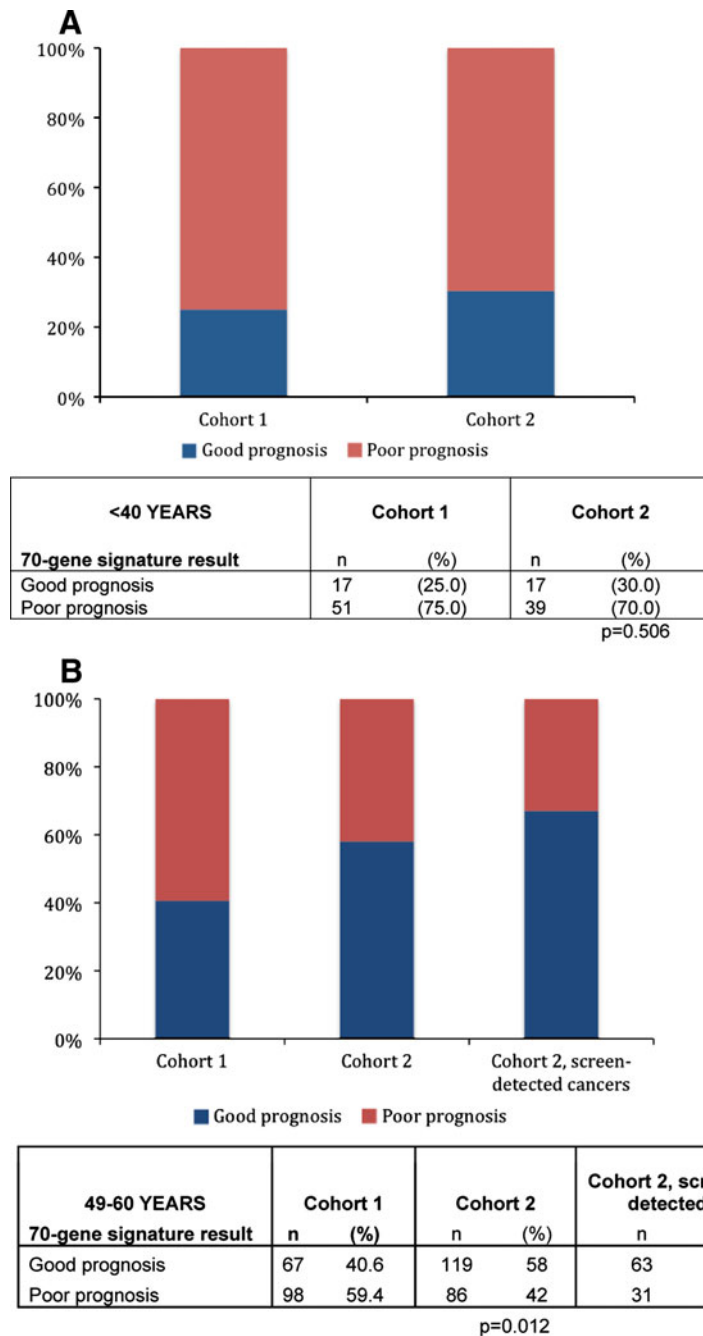
index scores was compared between the breast cancers from the Cohort 1, Cohort 2, and Cohort 2<sub>SD</sub> (Fig. 3). This analysis was limited to women aged 49–60 years, since 60 was the upper age limit for inclusion in the RASTER trial. The data show a significant shift towards a higher 70-gene index score in tumors in women from Cohort 2 and particularly in Cohort 2<sub>SD</sub> (Fig. 3a, bottom panel), compared with Cohort 1 (Fig. 3a, top panel). The median index score was 0.29, 0.48, and 0.51 in the women from Cohort 1, Cohort 2 (all), and Cohort 2<sub>SD</sub>. The distributions in the groups differed significantly (Mann–Whitney  $U = 3,271$ ,  $P < 0.01$ ). In Cohort 1, ages 49–60, 40.6% of cancers had index score greater than 0.4 (70-gene low risk threshold), compared with 58 and 67% in Cohort 2 and Cohort 2<sub>SD</sub>, respectively. In Cohort 1, 11.9% of cancers had an index score greater than 0.6 (ultralow threshold), compared with 31.7 and 31.1% in Cohort 2 and Cohort 2<sub>SD</sub>, respectively. Figure 3b shows the proportions of cancers that fall into the ultralow, low-non-ultralow, and high risk subsets. In Cohort 2, 67% of women had a low risk biology, almost half of which are ultralow risk. That compares with 40.6% with low risk signatures, less than a quarter of which are ultralow risk tumors, for the same age group in Cohort 1.

## Discussion

Molecular profiling is a tool that allows us to interrogate tumor biology. In this study, we used the 70-gene prognosis signature, MammaPrint<sup>TM</sup>, an FDA approved, robust gene array test, to investigate the biology of tumors 20 years ago, before the use of routine screening, and 5 years ago, after the introduction of population-based mammographic screening for breast cancer. We used tumor samples from the retrospective validation studies of patients who were diagnosed before population-based screening and samples from a prospective national demonstration project after the advent of screening. European countries have been very deliberate about the implementation of screening, so we can be confident that before national adoption and public financing, screening rates were low (less than 25% of the population was screened), but once screening was introduced through organized and publically financed programs, screening rates reached 75–80% of the population. The combination of access to the European and Dutch validation studies and the recent demonstration project using the 70-gene signature provided a unique opportunity to construct and compare the biology from a cohort from over 20 years ago, to a contemporary cohort. Screening was not routine in the first cohort but was in the second.

Several important observations can be made. First, the proportion of poor prognosis tumors varies significantly by age, with an increase in the likelihood of having a good

**Fig. 2** Mammographic screening results in an increase in the proportion of good prognosis cancers in Cohort 2, among women invited for population-wide screening. **a** Percentages of good versus poor prognosis cancers as a fraction of all cancers from Cohorts 1 and 2, respectively, are shown for patients under 40 years. There is no difference in the proportion of good prognosis cancers between Cohorts 1 and 2. Women in this age range did not undergo mammographic screening in either Cohort. **b** The percentages of good versus poor prognosis cancers as a fraction of all cancers in Cohort 1 and Cohort 2, respectively, are shown for patients aged 49–60 years. This age group was invited to participate in mammographic screening in Cohort 2, but not in Cohort 1. The third column shows the percentages of good versus poor prognosis cancers in the subset of Cohort 2 whose cancers were screen-detected. The *P* value refers the proportion of the low risk cancers as compared with Cohort 1

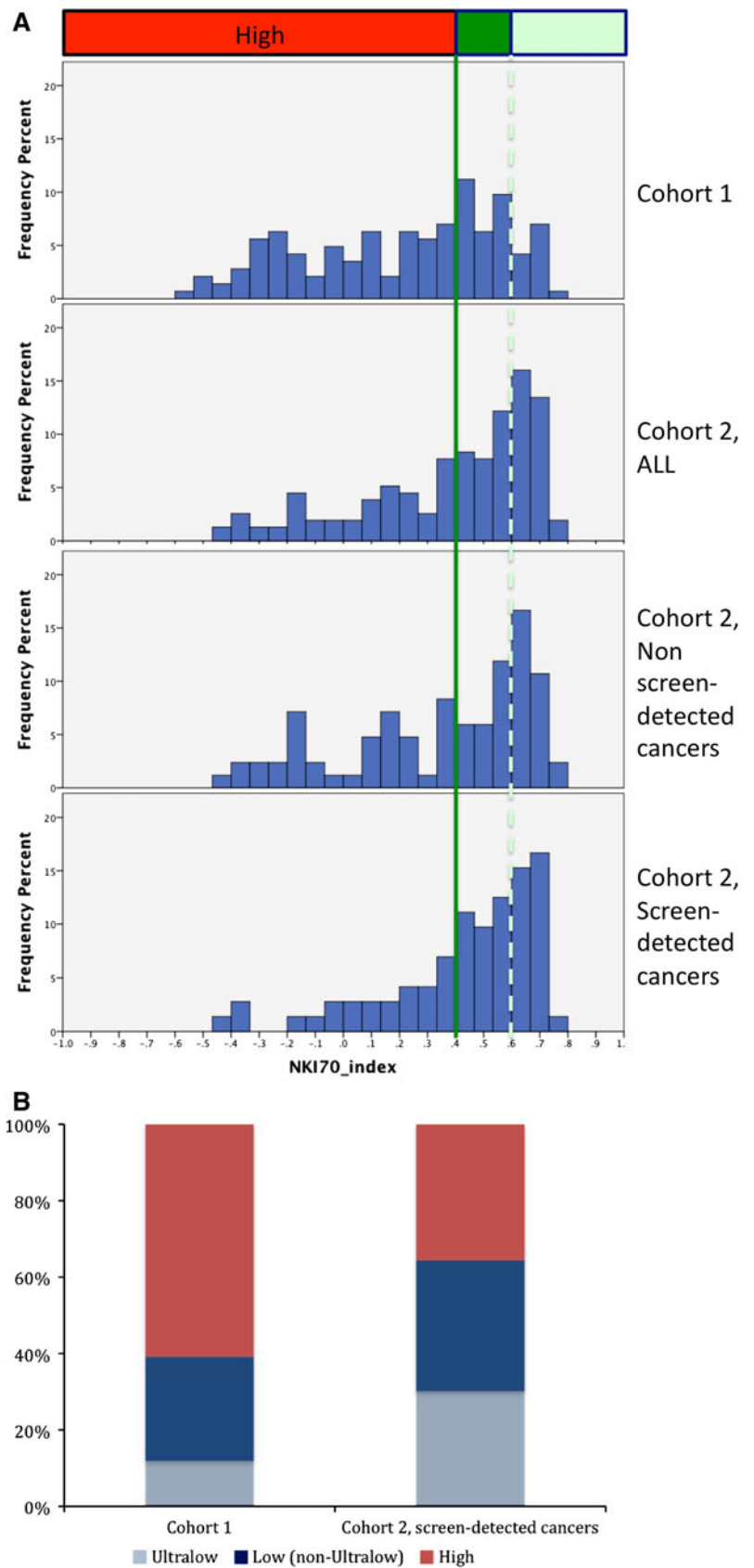


prognosis tumor as a woman ages. This is true for the combined population as well as for each individual cohort. Over 70% of tumors in women under the age of 40 are poor prognosis signature and this proportion has remained constant over the past 20 years. Interestingly, although tumors were smaller in younger women in the modern era (Cohort 2), the biology, as reflected by grade and 70-gene status did not change. In women 40 and over, there was a greater chance of having a good prognosis tumor in Cohort 2. The difference is larger in the age group 49–60, and largest for the women whose tumors were screen-detected. There is a corresponding shift to more favorable

clinicopathologic features in tumors as well, underscoring the association between clinicopathologic features and molecular profiles, and is consistent with the molecular data. It is likely that greater awareness about breast cancer is responsible for the detection of smaller tumors, even in the non-screened age groups. For women over 40 years of age, it is likely that there are factors in the population that may have changed over time. For women aged 49–60, the data suggest that a greater proportion of good prognosis tumors will be detected by screening, if they are present.

The histogram of tumors by 70-gene index score (Fig. 3) shows a significant shift to the right in Cohort 2, compared

**Fig. 3** Patients from Cohort 2 have tumors with a much higher proportion of low and ultralow risk biology. **a** Distribution of 70-gene prognosis index scores in women aged 49–60 years in the Cohort 1 (*top panel*), Cohort 2 (*second panel*), and the subsets of women from Cohort 2 with non-screen-detected (*third panel*) and screen-detected (*fourth panel*) cancers by frequency percent. An index score greater than 0.4 (*solid line*) corresponds to tumors with a good prognosis (low risk), and an index  $>0.6$  (*dashed line*) corresponds to ultralow risk. **b** Distribution of 70-gene signature risk groups as a percentage of total cancers in Cohort 1 versus screen-detected cancers from Cohort 2 in patients aged 49–60 years. The ultralow risk group is defined as index score  $>0.6$ , low risk (non-ultralow) is index score between 0.4 and 0.6, and high risk is index score  $<0.4$ . In the screen-detected group, 64% are low risk, approximately half of which are ultralow risk



with Cohort 1, especially for those women undergoing screening. In particular, the fraction of tumors with an index greater than 0.6 (designated as ultralow) is increased 200%. The significant increase in this fraction of the lowest risk tumors does indeed corroborate the notion that we maybe detecting, today, some tumors that might not come to clinical attention in the absence of screening. Interestingly, the distribution of grade in Cohort 2 is very similar to the distribution of the grade in the tumors detected in the Women's Health Initiative, where 78–82% had yearly mammography. (CHLEB 2003 JAMA), suggesting that Cohort 2 is representative of other cohorts. Compared with Cohort 1, where there is a 10% chance of finding an ultralow cancer, there is a 30% chance of finding an ultralow tumor in Cohort 2 if tumors are screen-detected. Interestingly, Welch and Black [2] recently estimated that 20% of detected breast cancers could represent "overdiagnosis".

An alternate explanation for this data is that the biology of tumors shifts in a given individual over time and that the ultralow risk tumors, if left intact and not found by screening, would progressively migrate towards a poorer prognosis signature. Evidence suggests that tumor biology does not change over time. Tumors appear to maintain the integrity of their molecular profile through treatment and recurrence [4, 10, 14]. Gene expression profiles of primary tumors are comparable with their distant metastases even if the metastatic disease appears after a long interval up to 15 years [10]. This was found to hold true for both intrinsic tumor type [14] as well as the 70-gene prognosis signature. Patients likely develop specific biologic tumor types, with different potential for metastasis and our findings suggest that screening enables the detection of very low risk tumors.

The data in this study clearly shows that, as women age, the likelihood of detection of good prognosis tumors rises substantially. The low risk and, in particular, the ultralow risk tumors are most always very endocrine sensitive. These findings help explain why older women with HR positive tumors have extremely good outcomes. CALGB 9343 [15] was a randomized trial for women over 70 who were randomized to hormone therapy alone or radiation. The incidence of distant metastases at 12 years median follow-up was only 3% in either arm.

The findings of increasing proportions of low risk tumors by age is important for informing screening policy and should provide critical input for informing screening intervals. With age, the biology of tumors shifts to lower risk lesions and slower growth fractions, making 2 year intervals reasonable. The RASTER trial was a population-based cohort, and thus the information about the biology of the tumors is likely to reflect the biology of tumors seen in other screening programs. The screening interval recommended in the Netherlands is every 2 years, and the majority of screen-detected cancers are low risk. One way

to inform the screening debate is to compare the types of tumors detected with annual and biannual screening to determine if there would be a projected benefit to more frequent screening. At least for the cohort in the RASTER trial, the majority of the tumors are good prognosis in the women aged 49–60 and more frequent screening would not necessarily improve outcomes [16, 17]. For women with screen-detected tumors, the ability to perform molecular tests and confirm the good prognostic nature of the tumor should give clinicians the confidence to pursue less aggressive interventions.

The data do not exclude the possibility that factors other than screening contribute to the shift to lower risk tumors in women over 40. The most likely is a shift in the population related to internal hormonal environment such as onset of menses, less and later age of child bearing, and increased use of alcohol that might also promote an increase in low risk tumors as has been shown in other areas of the world. Hormone replacement therapy is not likely to be a significant factor as the RASTER trial was initiated after the publication of the Women's Health Initiative study [18] that showed the link between combined hormone replacement therapy and increased risk for breast cancer [18], and prompted a precipitous fall off in use of HRT [19] world wide. Such effect was minimal in the Netherlands, however, as HRT use was already low in 2001 at 5.6% before the announcement of the WHI results, although use declined to 2.4% after publication by 2004 [20].

The retrospective nature of the datasets could have introduced bias. Some factors, however, minimize the chance of bias. The first is that we restricted the analysis to node negative cases only. Second, the rate of women getting mammograms may have been slightly higher than the 20% we estimated in Cohort 1 as the Karolinska Institute was screening women at increased risk before 1992, although the number of cases contributed from Sweden was small (data not shown). Finally, the fraction of low risk and ultralow risk tumors might even be under represented in the RASTER screening era cohort, as at least 7% of tumors were not profiled due to sampling failure [9] and the observation that T1 tumors have a higher proportion of low risk cancers as compared with T2 [21]. The 6 mm punch biopsy used at the time to collect the frozen tumor sample would minimize the ability to collect frozen material from the smallest tumors, thus potentially leading to an underrepresentation of low and ultralow risk tumors. On the other hand, some factors also increase the chance of finding features of poor prognosis biology in Cohort 1. Tumors from the NKI and other European centers might have been higher (disproportional) in severity given the fact that they were collected from referral centers. However, in Europe, there is less competition from community hospitals, more regional referral, and cases used were consecutive node negative tumors.



The significant shift in distribution towards especially favorable 70-gene prognosis MammaPrint index scores, provides the first molecular evidence of the increased detection of very low risk lesions over time. Screening appears to preferentially identify the low risk lesions in the population today. Given the extremely low risk for early recurrence carried by the ultralow risk signature, we have an example of how we might quantitatively apply the term, Indolent Lesions of Epithelial origin (IDLE) tumors, put forward in “Rethinking screening for Breast Cancer and Prostate Cancer” [1, 22]. Node negative tumors that have a 70-gene prognosis signature index of higher than 0.6 would qualify as IDLE tumors. Recurrences in this patient population would be predicted to be very infrequent, recurrences would be predicted to be late and likely controllable. It has long been known that women who recur after 10 years have disease that is much more indolent, a fact that is reflected by the excellent overall survival in the 70-gene good prognosis tumors [6, 8]. We are currently planning a validation study in a US based cohort, to determine the fraction of ultralow risk tumors in a screened population ([www.athenacarenetwork.org/](http://www.athenacarenetwork.org/)), but the information we have presented provide a rationale for integrating molecular profiling at the time of screening to help identify low risk tumors.

This study provides information that will allow us to improve screening. The data suggest that there is an opportunity to improve care by using validated predictors of risk for women with screen-detected tumors. Understanding that mammography preferentially detects slow to moderate growth tumors should be helpful on many fronts. Not only can this help us to guide the use of risk stratifying tools to avoid overtreatment, but it should also enable us to reset thresholds for biopsy for very low risk mammographic lesions (BIRADS 4A). The more indolent nature of the disease detected should give confidence to mammographers and surgeons to explore and test alternatives to biopsy of very low risk mammographic findings, which almost always turn out to be benign [23].

The observation that a substantial fraction of screen-detected cancers have low and ultralow risk is valuable information. These types of cancers may account for the cases that others consider “overdiagnosis” [24]. However, when we initiate screening, we do not know which women are likely to develop ultralow risk or IDLE tumors. We can, however, recognize that such tumors are commonly identified today, discuss this with our patients, and perform tests that elucidate the underlying biology of the tumors detected. We can use this information to guide treatment recommendations and as the basis for the development of clinical trials that test the safety of less aggressive treatments for patients with the lowest risk tumors.

**Acknowledgments** This study was supported by 5U01CA111234 (Early Detection Research Network), the UCSF Dean’s Medical Student Summer Research Fellowship (YS) and the Dutch Genomics Initiative ‘Cancer Genomics Center’. The Dutch Association of Comprehensive Cancer Centers and Michiel Schaapveld are acknowledged for providing stage specific incidence data. The authors thank Karen Sepucha for editing the manuscript.

## References

- Esserman L, Shieh Y, Thompson I (2009) Rethinking screening for breast cancer and prostate cancer. *JAMA* 302(15):1685–1692
- Welch HG, Black WC (2010) Overdiagnosis in cancer. *J Natl Cancer Inst* 102(9):605–613
- Ravdin PM et al (2001) Computer program to assist in making decisions about adjuvant therapy for women with early breast cancer. *J Clin Oncol* 19(4):980–991
- Sorlie T et al (2001) Gene expression patterns of breast carcinomas distinguish tumor subclasses with clinical implications. *Proc Natl Acad Sci USA* 98(19):10869–10874
- Paik S et al (2004) A multigene assay to predict recurrence of tamoxifen-treated, node-negative breast cancer. *N Engl J Med* 351(27):2817–2826
- van de Vijver MJ et al (2002) A gene-expression signature as a predictor of survival in breast cancer. *N Engl J Med* 347(25):1999–2009
- van’t Veer LJ et al (2002) Gene expression profiling predicts clinical outcome of breast cancer. *Nature* 415(6871):530–536
- Buyse M et al (2006) Validation and clinical utility of a 70-gene prognostic signature for women with node-negative breast cancer. *J Natl Cancer Inst* 98(17):1183–1192
- Bueno-de-Mesquita JM et al (2007) Use of 70-gene signature to predict prognosis of patients with node-negative breast cancer: a prospective community-based feasibility study (RASTER). *Lancet Oncol* 8(12):1079–1087
- Weigelt B et al (2003) Gene expression profiles of primary breast tumors maintained in distant metastases. *Proc Natl Acad Sci USA* 100(26):15901–15905
- Knauer M, Wenzl E, Rutgers EJ (2009) Gene expression profiling in breast cancer—design of a pooled database to address open questions. *Eur Surg* 41(5):221–227
- Mook S et al (2009) The 70-gene prognosis signature predicts early metastasis in breast cancer patients between 55 and 70 years of age. *Ann Oncol* 21:717–722
- Otten JD et al (2008) Impressive time-related influence of the Dutch screening programme on breast cancer incidence and mortality, 1975–2006. *Int J Cancer* 123(8):1929–1934
- Weigelt B et al (2005) Molecular portraits and 70-gene prognosis signature are preserved throughout the metastatic process of breast cancer. *Cancer Res* 65(20):9155–9158
- Hughes KS et al (2004) Lumpectomy plus tamoxifen with or without irradiation in women 70 years of age or older with early breast cancer. *N Engl J Med* 351(10):971–977
- Sihto H et al (2008) Molecular subtypes of breast cancers detected in mammography screening and outside of screening. *Clin Cancer Res* 14(13):4103–4110
- Berry DA et al (2009) Flawed inferences about screening mammography’s benefit based on observational data. *J Clin Oncol* 27(4):639–640; author reply 641–642
- Chlebowski RT et al (2003) Influence of estrogen plus progestin on breast cancer and mammography in healthy postmenopausal women: the Women’s Health Initiative Randomized Trial. *JAMA* 289(24):3243–3253

19. Ravdin PM et al (2007) The decrease in breast-cancer incidence in 2003 in the United States. *N Engl J Med* 356(16):1670–1674
20. de Jong-van den Berg LTW, Faber A, van den Berg PB (2006) HRT use in 2001 and 2004 in The Netherlands—a world of difference. *Maturitas* 54(2):193–197
21. Mook S et al (2010) Metastatic potential of T1 breast cancer can be predicted by the 70-gene MammaPrint signature. *Ann Surg Oncol* 17:1406–1413
22. Berry DA (2008) The screening mammography paradox: better when found, perhaps better not to find. *Br J Cancer* 98(11):1729–1730
23. Esserman L, Thompson I (2010) Solving the over diagnosis dilemma. *J Natl Cancer Inst* 102(9):582–583
24. Zahl PH, Maehlen J, Welch HG (2008) The natural history of invasive breast cancers detected by screening mammography. *Arch Intern Med* 168(21):2311–2316