Molecular Subtypes of Breast Cancers Detected in Mammography Screening and Outside of Screening

Harri Sihto,[1] Johan Lundin,[2,4] Tiina Lehtimäki,[2] Maarit Sarlomo-Rikala,[5] Ralf Bützow,[3,5] Kaija Holli,[6] Liisa Sailas,[8] Vesa Kataja,[8,9] Mikael Lundin,[2] Taina Turpeenniemi-Hujanen,[10] Jorma Isola,[7] Päivi Heikkilä,[5] and Heikki Joensuu[1,2]

Abstract

Purpose: The frequency and significance of gene expression profile-derived molecular subtypes of breast cancers found in mammography screening are unknown.

Experimental Design: We identified breast cancers diagnosed in women of any age living in defined geographic regions in Finland in 1991 to 1992 and collected clinical and pathologic data. Surrogates for the molecular subtypes were determined for 247 cancers found in organized mammography screening and 989 cancers detected outside of screening using immunohistochemistry or *in situ* hybridization. Molecular subtypes were defined as luminal A [estrogen receptor (ER) positive and/or progesterone receptor (PR) positive, *HER2*-], luminal B (ER+ and/or PR+, *HER2*+), basal-like (ER-, PR-, *HER2*-, cytokeratin 5+, and/or HER1+), *HER2*+/ER- (ER-, PR-, and *HER2*+), and unclassified. The median follow-up time was 9.4 years.

Results: The luminal type A was common (73.7%) and the *HER2*+/ER- type is rare (5.7%) in screen-detected cancer, and only 16% were *HER2* positive. Women with cancer diagnosed in screening at ages 50 to 69 years had similar molecular subtype distribution as women whose cancer was found outside of screening at age >69 years. In a multivariate model, cancer detection at screening independently predicted favorable distant disease-free survival when the molecular subtype was included as a covariate in addition to age, histologic grade, and cancer size. Women with small ($pT_1N_0M_0$) *HER2*-positive cancer had similar outcome regardless of the method of detection.

Conclusions: Molecular subtype distribution of screen-detected breast cancer differs from that of cancers found outside of screening and accounts in part for the better outcome of screen-detected cancer.

Women with breast cancer detected in mammography screening generally have more favorable outcome than those whose cancer is found outside of screening (1–3). Compared with cancers found by the patient herself, cancers detected in screening are usually smaller in size and are usually better differentiated and contain less tumor necrosis, have lower mitotic counts, express less frequently TP53 or HER2, have higher estrogen receptor (ER) and progesterone receptor (PR) contents, and are less frequently associated with regional lymph node metastases (4–8).

The generally favorable prognosis of breast cancer found in mammography screening may be related to several biases. Cancers found in screening are detected earlier during their natural course than those found outside of screening (the lead-time bias). Slow-growing cancers are more likely detected by periodic screening than rapidly growing cancers, leading to enrichment of more indolent cancers in a cohort of cancers detected by screening (the length bias). Individuals who participate in screening may not be representative to the entire population to be screened (selection bias), and some cancers detected in mammography screening might not manifest during the lifespan of the screened individual leading to overdiagnosis (9, 10). Some of these biases may have been obviated by prospective screening trials using breast cancer as an endpoint. Two large studies found that breast cancer detection in mammography screening preserves independent prognostic value in a multivariate analysis that includes also cancer size (7, 8), suggesting that early detection (lead-time bias) does not fully explain the generally better prognosis associated with breast cancer detection in mammography screening.

Breast cancers can be classified into distinct molecular subtypes that are associated with varying outcomes based on gene expression profiles (11–15). The major subtypes include the luminal A, luminal B, basal-like, HER2+/ER-, and normal breast-like subtypes (16). These subtypes can be characterized also using immunohistochemistry (17). Two of the subtypes, luminal A and B, consist of ER-positive cancers that frequently coexpress also PR and GATA-3 (18). Basal-like cancers are characteristically negative for ER, PR, and HER2 expression but may express HER1 (epidermal growth factor receptor), KIT (CD117), and basal cytokeratin (CK) 5/6 and 17 and frequently harbor mutated *TP53* and *BRCA1* (16, 19). HER2 expression and *HER2* amplification are associated with the *HER2*+/ER- and luminal B subtypes. The normal breast-like subtype is characterized with expression of genes related to nonepithelial cell types and adipose tissue, strong expression of basal epithelial genes, and low expression of luminal genes (11).

The length bias, detection of biologically more indolent cancers by screening, has thus far not been studied in depth as a potential factor that might explain the effect of the method of cancer detection on prognosis. In our previous study, none of the cancer biological factors evaluated by us could explain the favorable outcome associated with screen-detected breast cancer (7). It is currently not known whether there are differences in the proportions of the major molecular subtypes of breast cancers detected in screening and outside of screening and whether such differences might influence outcome.

Materials and Methods

Subjects. Women diagnosed with breast cancer within five defined geographic areas in Finland in 1991 or 1992 were identified from the files of the Finnish Cancer Registry (n = 2,930; ref. 20). This cohort, which is the basis of the study, comprises 53% of all breast cancers diagnosed in Finland in 1991 and 1992 (n = 5,551). Clinical data, including data on distant relapse and survival, were extracted from hospital case records and the files of the Finnish Cancer Registry and Statistics Finland. An effort was made to obtain detailed clinical information on 50 characteristics, including the histologic type and grade of breast cancer, the number of metastatic and nonmetastatic nodes, primary tumor size, tumor ER and PR content, treatment details, and follow-up data. These data were not available in 274 (9%) of the 2,930 patients. Patients who fulfilled the inclusion criteria but who were not identified in original computer search because the place of residence was outside the specified regions (n = 186) were included into study. Thus, 2,842 patients were entered into nationwide FinProg database.[11]

Subjects diagnosed with ductal or lobular carcinoma *in situ* were excluded from the study as well as those who had distant metastases at the time of the diagnosis, bilateral breast cancer, or other malignancy than breast cancer in history, except basal cell carcinoma or cervical carcinoma *in situ* (Supplementary Fig. S1). A subject was also excluded when the method for cancer detection was not known or if no breast surgery was carried out. A single subject may have been excluded for one or more reasons. Finally, we excluded cases where one or more biological markers required for tumor subtype classification was missing (n = 747). The required biomarker and clinical data were both available in 1,236 (43.5%) of the 2,842 cases who filled the study entry criteria. Of these cancers, 247 (19.6%) were detected in organized population-based mammography screening and 989 (80.4%) were detected outside of screening. The median follow-up time of the subjects included in study (n = 1,236) was 9.4 years after the diagnosis.

Permission to use formalin-fixed, paraffin-embedded tissues for research purposes was provided by the Ministry of Social Affairs and Health, Finland (permission 123/08/97).

Preparation of tissue microarrays. Tissue microarrays were prepared from formalin-fixed, paraffin-embedded tumor samples as described elsewhere (7). Sections of 5 μm were cut and processed for immunohistochemistry and chromogenic *in situ* hybridization (CISH).

Immunohistochemistry and CISH. Tumors were classified into five subtypes based on immunohistochemistry and CISH results. Tumor subtypes were classified using five expression markers, HER1 (erbB1), HER2 (erbB2), CK5/6, ER, and PR as described elsewhere (17). The molecular breast cancer subtypes were defined as luminal A (ER+ and/or PR+, *HER2*-), luminal B (ER+ and/or PR+, *HER2*+), basal-like (ER-, PR-, *HER2*-, CK5/6+, and/or HER1+), *HER2*+/ER- (ER-, PR-, and *HER2*+), and nonexpressor type (negative for all five key classifiers). *HER2* amplification as determined with CISH was preferred to HER2 expression determined by immunohistochemistry when cancers were assigned to one of the five molecular subtypes.

Immunostainings for ER, PR, HER1, HER2, TP53, Ki-67, and cyclooxygenase-2 were carried out and protein expression was evaluated as described earlier (21–23). Immunostainings were done using a mouse monoclonal antibody 6F11 (Novocastra Laboratories; 1:500 dilution) to evaluate ER expression. To assess expression of PR, HER1, HER2, TP-53, Ki-67, and cyclooxygenase-2, the following antibodies were used: 312 (Novocastra Laboratories; 1:500 dilution), NCL-EGFR (Novocastra Laboratories; 1:500 dilution), CB11 (Novocastra Laboratories; 1:200 dilution), DO7 (Novocastra Laboratories; 1:500 dilution), MM-1 (Novocastra Laboratories; 1:1,000 dilution), and 160112 (Cayman Chemical; 1:200 dilution), respectively. Immunohistochemistry for KIT was done using a rabbit polyclonal antibody A 4502 (DAKOCytomation; 1:300 dilution) and scored as described elsewhere (24).

For staining of GATA-3, antigen retrieval was carried out in a water bath [98°C for 40 min in 10 mmol/L sodium citrate buffer (pH 6.0)], and immunostaining was done using a mouse monoclonal HG3-31 antibody (SC-268; Santa Cruz Biotechnology; 1:300 dilution). CK5/6 expression was evaluated with a mouse monoclonal antibody M7237 (DAKO; 1:25 dilution), CK18 expression was evaluated with a mouse monoclonal antibody NCL-CK18 (Novocastra Laboratories; 1:20 dilution), and α-smooth muscle actin was evaluated with a mouse monoclonal antibody M0581 (DAKO; 1:100 dilution). Stainings were graded as negative (-), low expression (≤10% of cells are positive), moderate expression (11-50% of cells are positive), or high expression (51-100% of cells are positive).

HER2 gene copy number was assessed with CISH as described elsewhere (7). *HER2* amplification was considered to be present when six or more signals were detected per nucleus in >50% of cancer cells or when large gene copy clusters were present.

Statistical analysis. Frequency tables were analyzed using the χ^2 test. Life tables were calculated according to the Kaplan-Meier method. Distant disease-free survival was calculated from the date of the diagnosis to the date of detection of metastases outside of the locoregional area or to the date of death from breast cancer, whichever occurred first. Patients who died from an intercurrent disease were censored on the date of death. Survival curves were compared with the log-rank test. Multivariate survival analyses were done with the Cox proportional hazards model, entering the following covariates: molecular subtype, method of tumor detection, tumor size in centimeters, number of metastatic lymph nodes, histologic grade, and age at diagnosis. The assumption of proportional hazards was ascertained by assessment of log minus log survival plots. All *P* values are two tailed.

Results

Characteristics of breast cancers found in screening and outside of screening. Because population screening was mainly carried out in the age group of 50 to 59 years, the age distribution of women whose breast cancer was detected at mammography

---

[11] http://www.finprog.org

screening differed markedly from that of women whose cancer was found outside of screening (Table 1). As expected, breast cancers found in screening were smaller than those found outside of screening (83% versus 54%, respectively, had tumor diameter ≤2.0 cm) and had less often axillary nodal metastases (23% versus 38%, respectively). Screen-detected cancers were also generally better differentiated than cancers found outside of screening (35% versus 21%, respectively, were well differentiated), were more often one of the special histologic types than cancers found outside of screening (14% versus 9%), were more often ER positive (78% versus 69%), and had less often high Ki-67 expression (29% versus 37%). Screen-detected cancers also tended to be more often PR positive ($P = 0.052$) than cancers found outside of mammography screening

Table 1. Patient and tumor characteristics

Characteristic	Setting of breast cancer detection		P
	Screening (*n* = 247), *n* (%)	Outside of screening (*n* = 989), *n* (%)	
Age at diagnosis (y)			
≤39	2 (1)	82 (8)	<0.0001
40-49	30 (12)	238 (24)	
50-59	158 (64)	150 (15)	
60-69	48 (20)	198 (20)	
≥70	9 (4)	321 (32)	
Primary tumor diameter (cm)			
≤0.5	12 (5)	8 (1)	<0.0001
0.6-1.0	71 (30)	123 (13)	
1.1-2.0	114 (48)	384 (40)	
2.1-5.0	41 (17)	395 (41)	
>5.0	1 (0)	47 (5)	
NA	8	32	
No. positive axillary nodes			
0	186 (77)	573 (62)	<0.0001
1-3	44 (18)	231 (25)	
4-9	9 (4)	89 (10)	
≥10	2 (1)	27 (3)	
NA	6	69	
Histologic grade			
1	69 (35)	155 (21)	<0.0001
2	94 (48)	354 (48)	
3	32 (16)	223 (30)	
NA	52	257	
Histologic type			
Ductal	185 (75)	759 (77)	0.030
Lobular	27 (11)	142 (14)	
Special	35 (14)	88 (9)	
NA	0	0	
ER expression			
Negative	54 (22)	287 (29)	0.021
Positive	189 (78)	681 (69)	
NA	6	21	
Ki-67 expression			
Low-moderate	151 (71)	567 (63)	0.035
High	63 (29)	335 (37)	
NA	33	87	
PR expression			
Negative	85 (37)	426 (44)	0.052
Positive	145 (63)	542 (56)	
NA	17	21	
HER2 amplification			
Negative	209 (85)	789 (80)	0.085
Positive	38 (15)	200 (20)	
NA	0	0	
HER2 expression			
Negative	195 (84)	766 (80)	0.22
Positive	38 (16)	190 (20)	
NA	18	40	
TP53 expression			
Negative-low	165 (83)	723 (81)	0.60
Positive	34 (17)	166 (19)	
NA	48	100	
Adjuvant systemic therapy			
Not given	184 (75)	553 (57)	<0.0001
Given	60 (25)	414 (43)	
NA	3	22	

and tended to harbor less often amplified *HER2* oncogene (P = 0.085). Adjuvant systemic therapy, usually consisting of tamoxifen in ER-positive cases and/or i.v. cyclophosphamide, methotrexate, and 5-fluorouracil was administered to 43% of women who had their cancer detected outside of screening but only to 25% of women whose cancer was found in mammography screening (P < 0.0001). This difference in management likely reflects the generally larger size and less favorable biological features associated with breast cancers found outside of mammography screening.

Frequency of the molecular subtypes in breast cancer and protein expression. Presence of *HER2* amplification and expression of ER, PR, HER1, and CK5/6 were used to classify the tumors into five biological subtypes (Supplementary Table S1). The majority (n = 844; 68.3%) were of the luminal A subtype and 9.5% of luminal B, 9.7% of *HER2*+/ER-, 7.9% of basal-like, and 4.5% of the nonexpressor (all five markers negative) subtypes. When HER2 expression in immunohistochemistry was used instead of *HER2* amplification for classifying the samples, 68.1% were classified into the luminal A subtype, 9.7% to luminal B, 8.9% to *HER2*+/ER-, 8.1% to basal-like, and 5.1% to the nonexpressor subtype. Of the 747 cases that could not be classified due to lack of data on one or more of the five classifier factors, 27% were found in mammography screening and 73% were found outside of screening.

In addition to the five classifier factors, we examined expression of eight other proteins (HER2, Ki-67, TP53, GATA-3, CK18, KIT, α-smooth muscle actin, and cyclooxygenase-2) in the series (Supplementary Table S1). In line with the *HER2*

gene copy number analysis, HER2 protein expression was more common in the *HER2*+/ER- and the luminal B subtypes than in the other subtypes. High Ki-67 expression was less frequent in the luminal A subtype than in the rest of the molecular subtypes (22% versus 50-75%; P < 0.0001), suggesting that breast cancers of the luminal A subtype have a slower cell proliferation rate than the rest of the subtypes. Similarly, cancers of the luminal type A only rarely expressed TP53 protein (8%), whereas TP53 expression was relatively frequent in the other subtypes (24-52%; P < 0.0001). GATA-3 expression was common in the luminal A and B subtypes (90% and 89%, respectively) but rare in the basal-like type (10%; P < 0.0001) and cancers of the basal type expressed rarely also CK18 (76%). KIT expression was more common in the basal subtype (30%) compared with the rest of the subtypes (5-10% positive; P < 0.0001), whereas α-smooth muscle actin expression was rare (1-11%) in all of the subtypes. Moderate to high cyclooxygenase-2 expression was less frequent in the luminal A subtype (33%) compared with the rest of the subtypes (50-64%; P < 0.0001).

Frequency of the molecular subtypes in breast cancer found in mammography screening. The distribution of the molecular subtypes differed in screen-detected breast cancer compared with cancers found outside of mammography screening (Supplementary Table S2). In the age group of 50 to 69 years where most screen-detected cancers were diagnosed, 73.3% of cancers detected in mammography screening were of the luminal A subtype compared with 63.8% of cancers found outside of mammography screening, whereas the *HER2*+/ER- subtype, in particular, was more common among cancers
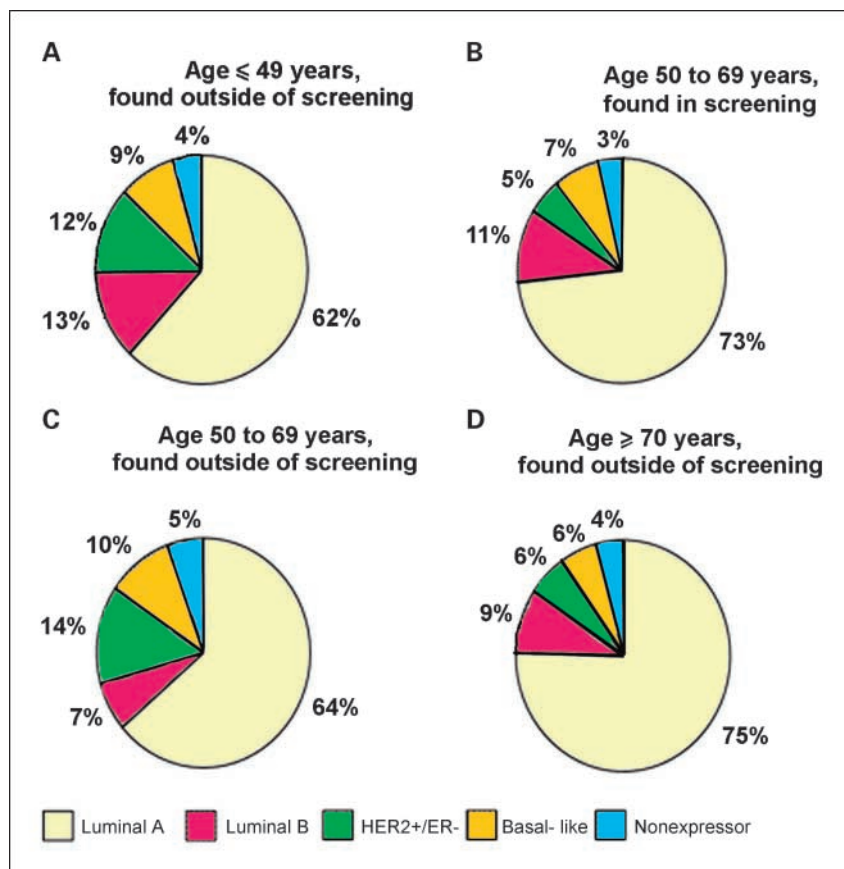

Fig. 1. Breast cancer molecular subtypes by age at diagnosis and the method of cancer detection. *A,* age ≤49 y, found outside of screening. *B,* age 50 to 69 y, found in screening. *C,* age 50-69 y, found outside of screening. *D,* age ≥70 y, found outside of screening.
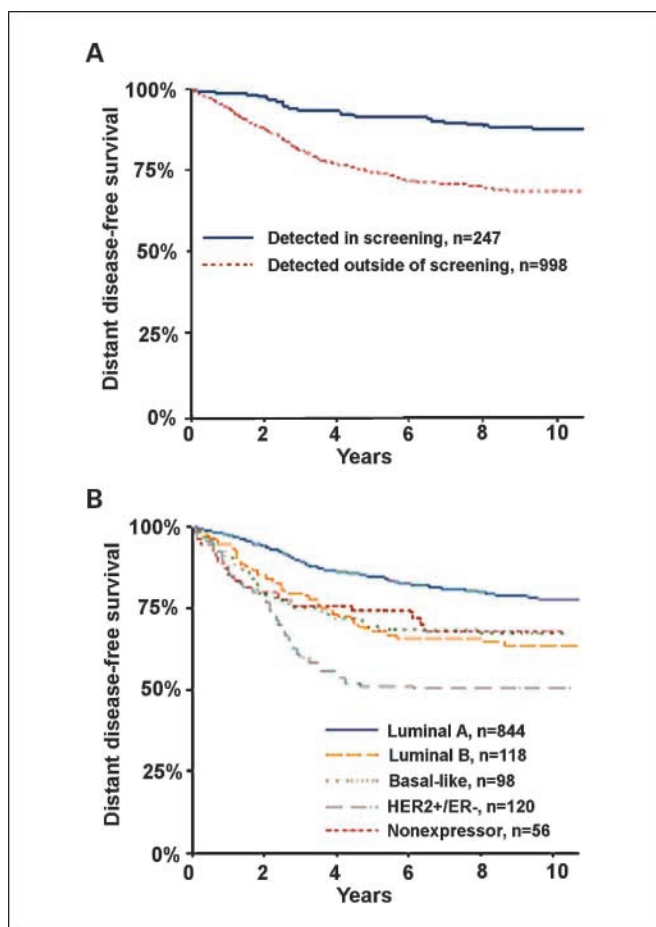
Fig. 2. Distant disease-free survival by the mode of detection (*A*) and the molecular subtype (*B*).

detected outside of screening than among screen-detected cancers (*P* = 0.049). Of note, the distribution of the molecular subtypes of cancers found in mammography screening in the age group of 50 to 69 years and that of breast cancers found outside of screening in the age group of ≥70 years did not differ significantly (*P* = 0.88; Fig. 1).

Distant disease-free survival in univariate analyses. Women with tumor detected by mammography had more favorable distant disease-free survival than those whose tumor was found outside of screening (*P* < 0.0001; Fig. 2A). The luminal subtype A was associated with the best distant disease-free survival and the *HER2*+/ER- subtype with the worst (*P* < 0.0001; Fig. 2B). Tumors detected in mammography screening had generally more favorable outcome than those detected outside of

screening irrespective of the molecular subtype (Table 2). In luminal type A, the 10-year distant disease-free survival was 75% among women whose cancer was found outside of screening and 90% when it was detected in mammography screening (*P* < 0.0001), and in *HER2*+/ER- subtype, the corresponding figures were 46% and 79%, respectively (*P* = 0.040).

The 10-year distant disease-free survival of women with *HER2*-positive cancer (identified by CISH) was 76% when cancer was detected in mammography screening and 53% when detected outside of screening (*P* = 0.011; Fig. 3A). In the subset of small (pT$_1$N$_0$M$_0$) *HER2* amplification-positive cancer, 74% of patients survived for 10 years free of distant metastases regardless of the mode of tumor detection (*P* = 1.0; Fig. 3B). In comparison, women diagnosed with *HER2*-negative, pT$_1$N$_0$M$_0$ cancer had significantly more favorable 10-year distant disease-free survival when cancer was detected in mammography screening than women whose cancer was found outside of screening (93% versus 86%; *P* = 0.0098; Fig. 3C).

Multivariate survival analysis. To assess whether cancer detection in mammography screening has independent influence on distant disease-free survival when cancer molecular subtype, size, histologic grade, and age at diagnosis are taken into account, we did a Cox multivariate survival analysis. In addition to histologic grade, tumor size, and number of axillary nodal metastases, both the molecular subtype and the cancer detection outside of mammography screening had independent influence on survival in this model (Table 3), whereas age at detection had no independent prognostic value.

We did another Cox multivariate analysis where we added adjuvant systemic treatment given (yes versus no) as a covariate in addition to the covariates listed in Table 3. This did not alter the results markedly. The same factors, including cancer molecular subtype (hazard ratio, 1.12; 95% confidence interval, 1.05-1.20; *P* = 0.001) and cancer detection outside of mammography screening (hazard ratio, 1.83; 95% confidence interval, 1.18-2.88; *P* = 0.008), remained as independent prognostic factors in the model, whereas administration of adjuvant systemic treatment did not have independent influence on distant disease-free survival (hazard ratio, 1.15; 95% confidence interval, 0.83-1.59; *P* = 0.40), suggesting that adjuvant systemic treatment administered has little influence on the prognostic value of the factors listed in Table 3. Similarly, cancer histologic type (ductal versus lobular versus special type) had no prognostic value when added as a covariate to the model.

Discussion

We compared the frequency of molecular subtypes in breast cancers detected in population-based mammography screening and outside of screening and evaluated whether variance in the

Table 2. Distant disease-free survival by the molecular subtype

Molecular type	n	10-y distant disease-free survival		P
		Found in screening (%)	Found outside of screening (%)	
Luminal A	844	90	75	<0.0001
Luminal B	118	75	61	0.21
HER2+/ER-	120	79	46	0.040
Basal	98	88	63	0.087
Nonexpressor	56	90	63	0.10

molecular subtype distribution might explain the generally more favorable prognosis associated with cancers detected in mammography screening. The luminal A subtype, which is generally associated with a favorable outcome (11, 16), was more frequent among screen-detected breast cancers than cancers detected outside of screening, whereas the *HER2*+/ER- subtype, which is generally associated with an unfavorable clinical course (11, 16, 17), was relatively rare in cancers
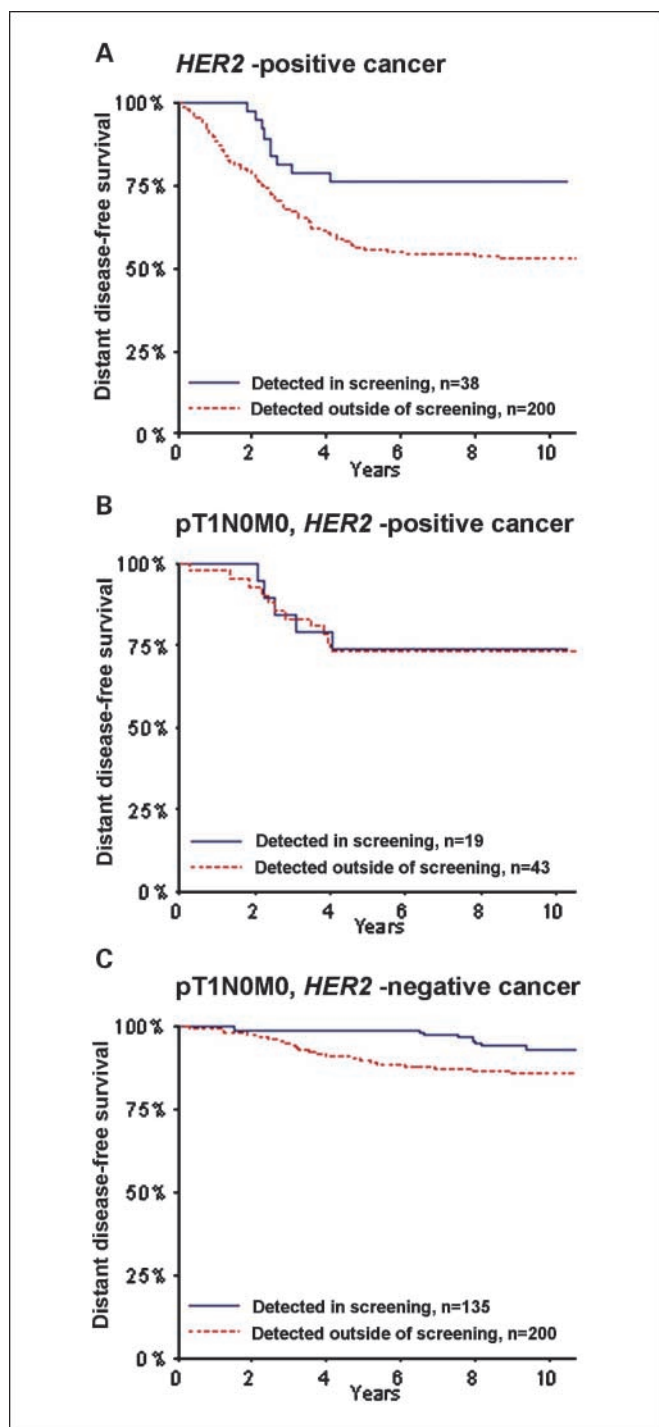
Fig. 3. Influence of the method of detection and HER2 status in distant disease-free survival. *A,* all subjects with *HER2*-positive breast cancer. *B,* subjects with $pT_1N_0M_0$ *HER2*-positive cancer. *C,* subjects with $pT_1N_0M_0$ *HER2*-negative cancer.

detected in mammography screening. However, in a multivariate analysis, cancer detection at screening remained an independent prognostic variable when the molecular subtype was included as a covariate in the model together with tumor size and other covariates, suggesting that the molecular subtype distribution can only partially explain the favorable outcome associated with cancer detection in mammography screening.

Interestingly, women whose cancer was diagnosed in screening at the age of 50 to 69 years had strikingly similar distribution of cancer molecular subtypes as women whose cancer was found outside of screening at the age of ≥70 years. This finding is compatible with a length bias (screening tends to detect slow-growing, less aggressive cancers) and also with a lead-time bias (screening tends to detect cancers earlier during their natural course). Because the mean sojourn time (the mean preclinical detectable period) of breast cancer has been estimated to be ~4 years in the age group of 50 to 69 years (25–28), the length bias may be a major factor that accounts for this finding. Yet, there is little doubt that also aggressive cancers may be detected in screening; women with small ($pT_1N_0M_0$) *HER2*-positive node-negative cancers had similar outcome regardless of their method of detection. Of note, these cancers were associated with only 74% of 10-year distant disease-free survival despite the small primary tumor size.

To our knowledge, the frequency and outcome of the molecular subtypes of screen-detected breast cancer have not been evaluated earlier. In a small study where 95 interval cancers diagnosed in a Norwegian Breast Cancer Screening Program were compared with 95 screen-detected breast cancers, the interval cancers expressed more often CK5/6 (22% versus 10%, respectively) and P-cadherin (24% versus 10%) than breast cancers found in screening, suggesting that the basal type is more frequent among cancers detected outside of screening (29). The portion of *HER2*-positive cancer has also been found to be higher in interval cancers than screen-detected cancers (5). Besides age at diagnosis and the method of cancer detection, the proportions of the molecular subtypes appear to depend on race. In one study, basal-like breast cancer was found to be considerably more frequent among premenopausal African American women (39%) than premenopausal non-African Americans (16%), postmenopausal African Americans (14%), or postmenopausal non-African American women (16%; ref. 17).

The proportion of basal and *HER2*+/ER- subtypes tends to be considerably higher in breast cancer patient series that consists of women treated with chemotherapy than among population-based series, suggesting selection of younger breast cancer patients whose cancer has been detected outside of mammography screening to chemotherapy trials (30, 31). Results from a few retrospective exploratory studies suggest that cancers of the basal cell or the *HER2*+/ER- subtype may respond more often to chemotherapy than cancers of the luminal type (32, 33), although women with cancer of the *HER2*+/ER- or the basal-like type tend to have an unfavorable outcome (29, 30).

The frequency of *HER2*-positive breast cancer is commonly cited to be 20% to 25% of all breast cancers (34, 35). These frequencies are similar to the ones we found among women whose cancer was detected outside of mammography screening at the age of <70 years. However, the proportion of *HER2*-positive breast cancer was only 16% among women whose cancer was detected in mammography screening at the age of 50 to 69 years and 15% among those whose cancer was found

Table 3. Cox multivariate analysis of distant disease-free survival

Covariate	Hazard ratio (95% confidence interval)	*P*
Molecular subtype		
Luminal A	1.00	
Basal-like	1.43 (0.82-2.51)	0.21
Nonexpressor	1.32 (0.83-2.11)	0.24
Luminal B	1.55 (1.02-2.35)	0.04
HER2+/ER-	1.88 (1.30-2.74)	0.001
Tumor size*	1.17 (1.07-1.27)	<0.0001
No. positive axillary lymph nodes [†]	1.14 (1.11-1.18)	<0.0001
Detection outside of screening	1.79 (1.14-2.81)	0.011
Histologic grade (grade 3 or 2 vs 1)	2.59 (1.56-4.27)	<0.0001
Age at diagnosis (y)		
35-69	1.00	
<35	0.93 (0.47-1.84)	0.84

*Hazard provided per 1 cm of the longest diameter of the tumor.
[†]Hazard provided per one metastatic node.

outside of screening at the age of ≥70 years. Hence, the proportion of *HER2*-positive breast cancers diagnosed may vary markedly depending on extent of organized mammography screening carried out in the population.

In addition to the five classifier markers, we analyzed also expression of eight other proteins in the series (Supplementary Table S1). In general, the results of these analyses are in line with those reported earlier. Expression of CK18, an intermediate filament, is associated particularly with the luminal cell types A and B, but we found CK18 expression also in almost all (98%) *HER2*+/ER- type tumors. GATA-3, a zinc-finger transcription factor that has a well-defined role in cell-fate specification in the immune system, in the kidney and other tissues, is a defining marker of the "luminal" subtypes of breast cancer (36). In the mammary gland, GATA-3 is expressed only by the epithelium, where its expression increases during early pregnancy and is associated with ER expression (18, 36–38). In line with this, most (90%) luminal cancers but only few (10%) basal-type breast cancers expressed GATA-3 in the present series. KIT tyrosine kinase was expressed in 30% of basal-type breast carcinomas as reported by others (19).

A limitation of the present study is that we could not analyze one or more of the five classifier markers from tumor tissue sample of 747 (37.7%) women of the 1,983 potentially eligible subjects. This was usually due to a lack of representative tissue, an unsuccessfully cut core from the donor block, or presence of technically unsuccessful staining. Twenty-six percent (*n* = 192) of these 747 cancers were found in screening compared with 20% (*n* = 247) of the 1,236 cases that were classified successfully. This difference is statistically significant (*P* = 0.015) and is likely explained by the generally small size of the screen-detected cancer that often results in a lack of tissue

available for analysis. The influence of a lack of adequate amount of starting material in the smallest tumors, if any, on the current results is not known.

Besides the molecular subtype, other factors that might explain the generally more favorable prognosis of cancers detected in screening compared with cancers of similar size detected outside of screening remain unknown. Breast tumors that cause pain, abnormal sensations, or bleeding might be more invasive or secrete more proinflammatory or cytotoxic metabolites than asymptomatic cancers detected in screening, or the stromal composition of self-detected cancers might be different resulting in a more easily detectable palpable lump. We did not evaluate cancer invasiveness or migration-associated factors or cell adhesion molecule expression in the present study.

We conclude that ∼75% of all breast cancers detected in population-based mammography screening at the age of 50 to 69 years are luminal type A breast cancers, a subtype that is generally associated with a favorable outcome. Women with cancer detected in mammography screening have relatively rarely *HER2*+/ER- or basal type of breast cancer, which subtypes are characterized by a more aggressive natural history than the luminal type A cancer. The molecular subtype distributions explain only in part the generally favorable outcome of screen-detected breast cancer even when cancer size is accounted for. The distribution of the molecular subtypes of breast cancers found in mammography screening at the age of 50 to 69 years resembles that of cancers detected outside of mammography screening in women ages >69 years.

Disclosure of Potential Conflicts of Interest

No potential conflicts of interest were disclosed.

References

1. Senie RT, Lesser M, Kinne DW, Rosen PP. Method of tumor detection influences disease-free survival of women with breast carcinoma. Cancer 1994;73: 1666–72.
2. Burrell HC, Sibbering DM, Wilson AR, et al. Screening interval breast cancers: mammographic features and prognostic factors. Radiology 1996;199:811–7.
3. Moody-Ayers SY, Wells CK, Feinstein AR. "Benign"

tumors and "early detection" in mammography-screened patients of natural cohort with breast cancer. Arch Intern Med 2000;160:1109–15.
4. Porter PL, El-Bastawissi AY, Mandelson MT, et al. Breast tumor characteristics as predictors of mammographic detection: comparison of interval- and screen-detected cancers. J Natl Cancer Inst 1999;91: 2020–8.

5. Crosier M, Scott D, Wilson RG, Griffiths CD, May FE, Westley BR. Differences in Ki67 and c-erbB2 expression between screen-detected and true interval breast cancers. Clin Cancer Res 1999;5:2682–8.
6. Groenendijk RP, Bult P, Tewarie L, et al. Screen-detected breast cancers have a lower mitotic activity index. Br J Cancer 2000;82:381–4.
7. Joensuu H, Lehtimäki T, Holli K, et al. Risk for distant

recurrence of breast cancer detected by mammography screening or other means. JAMA 2004;292: 1064–73.

8. Shen Y, Yang Y, Inoue LY, Munsell MF, Miller AB, Berry DA. Role of detection method in predicting breast cancer survival: analysis of randomized screening trials. J Natl Cancer Inst 2005;97:1195–203.

9. Kopans DB, Monsees B, Feig SA. Screening for cancer: when is it valid? Lessons from the mammography experience. Radiology 2003;229:319–27.

10. Zahl PH, Strand BH, Maehlen J. Incidence of breast cancer in Norway and Sweden during introduction of nationwide screening: prospective cohort study. BMJ 2004;328:921–4.

11. Sørlie T, Perou CM, Tibshirani R, et al. Gene expression patterns of breast carcinomas distinguish tumor subclasses with clinical implications. Proc Natl Acad Sci U S A 2001;98:10869–74.

12. van't Veer LJ, Dai H, van de Vijver MJ, et al. Gene expression profiling predicts clinical outcome of breast cancer. Nature 2002;415:530–6.

13. Kang YB, Siegel PM, Shu WP, et al. A multigenic program mediating breast cancer metastasis to bone. Cancer Cell 2003;3:537–49.

14. Wang Y, Klijn JG, Zhang Y, et al. Gene-expression profiles to predict distant metastasis of lymph-node-negative primary breast cancer. Lancet 2005;365: 671–9.

15. Naume B, Zhao X, Synnestvedt M, et al. Presence of bone marrow micrometastasis is associated with different recurrence risk within molecular subtypes of breast cancer. Mol Oncol 2007;1:160–71.

16. Sorlie T, Tibshirani R, Parker J, et al. Repeated observation of breast tumor subtypes in independent gene expression data sets. Proc Natl Acad Sci U S A 2003; 100:8418–23.

17. Carey LA, Perou CM, Livasy CA, et al. Race, breast cancer subtypes, and survival in the Carolina Breast Cancer Study. JAMA 2006;295:2492–502.

18. Hoch RV, Thompson DA, Baker RJ, Weigel RJ. GATA-3 is expressed in association with estrogen receptor in breast cancer. Int J Cancer 1999;84:122–8.

19. Laakso M, Tanner M, Nilsson J, et al. Basoluminal carcinoma: a new biologically and prognostically distinct entity between basal and luminal breast cancer. Clin Cancer Res 2006;12:4185–91.

20. Lundin J, Lundin M, Holli K, et al. Omission of histologic grading from clinical decision making may result in overuse of adjuvant therapies in breast cancer: results from a nationwide study. J Clin Oncol 2001; 19:28–36.

21. Joensuu H, Isola J, Lundin M, et al. Amplification of erbB2 and erbB2 expression are superior to estrogen receptor status as risk factors for distant recurrence in $pT_1N_0M_0$ breast cancer: a nationwide population-based study. Clin Cancer Res 2003;9:923–30.

22. Ristimäki A, Sivula A, Lundin J, et al. Prognostic significance of elevated cyclooxygenase-2 expression in breast cancer. Cancer Res 2002;62:632–5.

23. Lassus H, Sihto H, Leminen A, et al. Gene amplification, mutation, and protein expression of EGFR and mutations of ERBB2 in serous ovarian carcinoma. J Mol Med 2006;84:671–81.

24. Sihto H, Tynninen O, Bützow R, Saarialho-Kere U, Joensuu H. Endothelial cell KIT expression in human tumors. J Pathol 2007;211:481–8.

25. Brekelmans CT, Westers P, Faber JA, Peeters PH, Collette HJ. Age specific sensitivity and sojourn time in a breast cancer screening programme (DOM) in The Netherlands: a comparison of different methods. J Epidemiol Community Health 1996;50:68–71.

26. Paci E, Duffy SW. Modelling the analysis of breast cancer screening programmes: sensitivity, lead time and predictive values in the Florence district programme (1975-1986). Int J Epidemiol 1991;20:852–8.

27. Boer R, de Koning H, Threlfall A, et al. Cost effectiveness of shortening screening interval or extending age range of NHS breast screening programme: computer simulation study. BMJ 1998;317:376–9.

28. Boer R, de Koning HJ, van der Maas PJ. A longer breast carcinoma screening interval for women age older than 65 years? Cancer 1999;86:1506–10.

29. Collett K, Stefansson IM, Eide J, et al. A basal epithelial phenotype is more frequent in interval breast cancers compared with screen detected tumors. Cancer Epidemiol Biomarkers Prev 2005;14:1108–12.

30. Diallo-Danebrock R, Ting E, Gluz O, et al. Protein expression profiling in high-risk breast cancer patients treated with high-dose or conventional dose-dense chemotherapy. Clin Cancer Res 2007;13:488–97.

31. Hannemann J, Kristel P, van Tinteren H, et al. Molecular subtypes of breast cancer and amplification of topoisomerase IIα: predictive role in dose intensive adjuvant chemotherapy. Br J Cancer 2006;95:1334–41.

32. Carey LA, Dees EC, Sawyer L, et al. The triple negative paradox: primary tumor chemosensitivity of breast cancer subtypes. Clin Cancer Res 2007;13:2329–34.

33. Conforti R, Boulet T, Tomasic G, et al. Breast cancer molecular subclassification and estrogen receptor expression to efficacy of adjuvant anthracycline-based chemotherapy: a biomarker study from two randomized trials. Ann Oncol 2007;18:1477–83.

34. Pauletti G, Dandekar S, Rong H, et al. Assessment of methods for tissue-based detection of the HER-2/ *neu* alteration in human breast cancer: a direct comparison of fluorescence *in situ* hybridization and immunohistochemistry. J Clin Oncol 2000;18:3651–64.

35. Borg A, Tandon AK, Sigurdsson H, et al. HER-2/*neu* amplification predicts poor survival in node-positive breast cancer. Cancer Res 1990;50:4332–7.

36. Asselin-Labat ML, Sutherland KD, Barker H, et al. Gata-3 is an essential regulator of mammary-gland morphogenesis and luminal-cell differentiation. Nat Cell Biol 2007;9:201–9.

37. Rakha EA, Putti TC, Abd El-Rehim DM, et al. Morphological and immunophenotypic analysis of breast carcinomas with basal and myoepithelial differentiation. J Pathol 2006;208:495–506.

38. Naylor MJ, Ormandy CJ. Gata-3 and mammary cell fate. Breast Cancer Res 2007;9:302.

Clinical Cancer Research

Molecular Subtypes of Breast Cancers Detected in Mammography Screening and Outside of Screening

Harri Sihto, Johan Lundin, Tiina Lehtimäki, et al.

Updated version	Access the most recent version of this article at: http://clincancerres.aacrjournals.org/content/14/13/4103
Supplementary Material	Access the most recent supplemental material at: http://clincancerres.aacrjournals.org/content/suppl/2008/09/03/14.13.4103.DC1

Cited articles	This article cites 38 articles, 15 of which you can access for free at: http://clincancerres.aacrjournals.org/content/14/13/4103.full#ref-list-1
Citing articles	This article has been cited by 13 HighWire-hosted articles. Access the articles at: http://clincancerres.aacrjournals.org/content/14/13/4103.full#related-urls

E-mail alerts	Sign up to receive free email-alerts related to this article or journal.
Reprints and Subscriptions	To order reprints of this article or to subscribe to the journal, contact the AACR Publications Department at pubs@aacr.org.
Permissions	To request permission to re-use all or part of this article, use this link http://clincancerres.aacrjournals.org/content/14/13/4103. Click on "Request Permissions" which will take you to the Copyright Clearance Center's (CCC) Rightslink site.