VOLUME 34 · NUMBER 3 · JANUARY 20, 2016

JOURNAL OF CLINICAL ONCOLOGY

ORIGINAL REPORT

Associations of Breast Cancer Risk Prediction Tools With Tumor Characteristics and Metastasis

Johanna Holm, Jingmei Li, Hatef Darabi, Martin Eklund, Mikael Eriksson, Keith Humphreys, Per Hall, and Kamila Czene

All authors: Karolinska Institutet, Stockholm, Sweden

Published online ahead of print at www.jco.org on November 30, 2015.

Supported by the Swedish Research Council (Grant No. 2014-2271), Swedish Cancer Society (Grant No. CAN 2013/469), Cancer Society in Stockholm (Grant No. 141092), and the Cancer Risk Prediction Center (CRisP; www.crispcenter.org), a Linneus Centre (Contract ID 70867902) financed by the Swedish Research Council. K.H. is supported by the Swedish Research Council (Grant No. 521-2011-3205) and the Swedish Cancer Society (contract ID 110600). J.L. is a UNESCO-L'Oréal International Fellow.

Terms in blue are defined in the glossary, found at the end of this article and online at www.jco.org.

Authors' disclosures of potential conflicts of interest are found in the article online at www.jco.org. Author contributions are found at the end of this article.

Corresponding author: Johanna Holm, Karolinska Institutet, Department of Medical Epidemiology & Biostatistics, Nobels väg 12 A, SE-171 77 Solna, Sweden; e-mail: johanna.holm@ki.se.

0732-183X/16/3403w-251w/$20.00

DOI: 10.1200/JCO.2015.63.0624

ABSTRACT

Purpose
The association between established risk factors for breast cancer and subtypes or prognosis of the disease is not well known. We analyzed whether the Tyrer-Cuzick–predicted 10-year breast cancer risk score (TCRS), mammographic density (MD), and a 77-single nucleotide polymorphism polygenic risk score (PRS) were associated with breast cancer tumor prognosticators and risk of distant metastasis.

Patients and Methods
We used a case-only design in a population-based cohort of 5,500 Swedish patients with breast cancer. Logistic and multinomial logistic regression of outcomes, estrogen receptor (ER) status, lymph node involvement, tumor size, and grade was performed with TCRS, PRS, and percent MD as exposures. Cox proportional hazard models were used to estimate hazard ratios (HRs) of distant metastasis.

Results
Women at high risk for breast cancer based on PRS and/or TCRS were significantly more likely to be diagnosed with favorable prognosticators, such as ER-positive and low-grade tumors. In contrast, PRS weighted on ER-negative disease was associated with ER-negative tumors. When stratifying by age, the associations of TCRS with favorable prognosticators were restricted to women younger than age 50. Women scoring high in both TCRS and PRS had a lower risk of distant metastasis (HR, 0.69; 95% CI, 0.49 to 0.98). MD was not associated with any of the examined prognosticators.

Conclusion
Women at high risk for breast cancer based on genetic and lifestyle factors were significantly more likely to be diagnosed with breast cancers with a favorable prognosis. Better knowledge of subtype-specific risk factors could be vital for the success of prevention programs aimed at lowering mortality.

J Clin Oncol 34:251-258. © 2015 by American Society of Clinical Oncology

INTRODUCTION

Breast cancer is a heterogeneous disease, with both environmental and genetic causes. Currently known risk factors can be used in prediction modeling to identify women at risk for disease, a necessary step for improving early detection, as well as offering preventative interventions.[1] The Tyrer-Cuzick model has been developed to identify women at increased risk for breast cancer, incorporating hormonal, lifestyle, and reproductive risk factors, together with family history, in the estimation of risk.[2] Additionally, large-scale genome-wide association studies have identified many low-risk common variant single-nucleotide polymorphisms (SNPs) that partially explain disease occurrence when analyzed as a polygenic risk score (PRS). Recently, a PRS based on 77-SNP variants was developed to predict breast cancer risk.[3] Another factor used for risk prediction is mammographic density (MD), one of the strongest risk factors for breast cancer.[4] These tools show future promise for determining breast cancer risk in healthy women, and their use has been suggested to improve performance when used in combinations.[5-8]

Identification of women at high risk for breast cancer is a pivotal part of successful clinical prevention aimed at lowering disease mortality. Importantly, if disease burden and mortality are to be lowered, it is necessary to be able to identify women at risk for all subtypes of the disease. However, the prediction tools mentioned earlier may be differentially associated with subtypes of

the disease, given previous findings of separate etiologies for distinct molecular subtypes of breast cancer.[9-11] We investigated whether Tyrer-Cuzick–predicted 10-year breast cancer risk score (TCRS), PRS, and percent MD values differed, on average, according to tumor prognosticators and metastasis in a large population-based cohort of breast cancer patients in Sweden.

PATIENTS AND METHODS

All female breast cancer patients younger than 80 years of age diagnosed between 2001 and 2008 in Stockholm-Gotland County, Sweden (n = 9,348) were invited to donate blood and answer a Web-based questionnaire. In all, 5,715 women (61%) consented and returned detailed information regarding lifestyle, hormonal, and reproductive factors, as well as consent to retrieve mammographic images. Women with a previous breast cancer diagnosis (n = 215) were excluded from the study, leaving 5,500 women for analysis.

Tumor and patient characteristics were collected for all women by links to the high-coverage Stockholm-Gotland Regional Breast Cancer Quality Register.[12] Lymph node involvement was dichotomized into positive or negative. Tumor size was categorized into less than 20 mm, 20 to 40 mm, and greater than 40 mm in diameter. Estrogen receptor (ER) status was recorded as negative or positive in the registers, determined by radioimmunoassay or immunohistochemistry. Registry information was essentially complete (98%) for tumor size and lymph node status, whereas missing data were higher for ER status (80% complete). Grade was available from 2004 onward, with 93% completeness.

The individual 10-year and 5-year risk of breast cancer was assessed at age of diagnosis with the Tyrer-Cuzick model, using the International Breast Cancer Intervention Study (IBIS) tool.[2] Information on reproductive, familial, anthropometric, and hormonal variables incorporated was available in a Web-based questionnaire, self-reported by the women on study recruitment. Information on history of benign breast disease and second-degree family history was not available in the data, yielding potential underestimation of the risk. Data were 95% to 100% complete on all used questionnaire variables.

Blood samples from participants were genotyped on a custom Illumina iSelect Array (iCOGS Array) consisting of 211,155 selected SNPs.[13] Missing genotypes were imputed using 1000 Genomes (phase I integrated variant set release [v3] in National Center for Biotechnology Information build 37 [hg19] coordinates). Full details of the construction of the PRS has been described previously.[14] The polygenic score for each patient was evaluated using the PLINK command (–score), using weights reported in Mavaddat et al.[3] Three separate PRSs were constructed, corresponding to weights from analyses on breast cancer overall (PRS), ER-negative disease (PRS ER-negative), and ER-positive disease (PRS ER-positive).

Analog mammographic images were retrospectively collected from five main radiology units. A prediagnostic analog film in the cancer-free breast was available for 3,488 women. All mammographic images were digitized with an Array 2905HD Laser Film Digitizer (Array, Tokyo, Japan). Percent MD was assessed with an automated algorithm developed and described by Li et al.[15] The algorithm is trained to distinguish dense breast areas from nondense areas by training on image segmentation data measured through Cumulus.[16] All study participants provided informed consent, and the ethical committee at Karolinska Institutet approved the study.

Statistical analysis was performed in R (version 3.1.0).[17] Each statistical test was two-sided, with an alpha level set at .05. Any women with missing information were excluded from analysis. Correlation coefficients between the scores were calculated using Spearman's rank method. The binary outcomes, ER status and lymph node involvement, were modeled with binomial logistic regression analysis. Multinomial logistic regression (using the R "nnet"' package) was used for the categorical outcomes of grade and tumor size. The three risk prediction tools, MD, PRS, and TCRS, were entered as standardized linear scores in all regression analyses. Each score was modeled separately to assess its individual associations with the outcomes. Analyses were performed overall, stratified by ER status, and stratified by women younger than age 50 and 50 years or older. For PRS, analysis was also done using PRS ER-negative and PRS ER-positive. All analyses concerning MD were adjusted for age. For sensitivity analyses, separate estimations were performed for women without a first-degree relative with breast cancer (defined as mother or sister). Analysis using TCRS 5-year estimates was also performed as a supplementary analysis (Appendix Table A1 and A2, online only).

Cox proportional hazard regression models (using the R "survival" package) were used to estimate hazard ratios (HRs) of distant metastasis for each risk prediction tool, with time since diagnosis as the underlying time scale. Subjects were followed from the date of diagnosis until the date of distant metastasis, date of death from any cause, or end of follow-up on September 12, 2014, whichever came first. Women with metastasis at the time of diagnosis were excluded (n = 21). Information on the date of metastasis was available from the Regional Breast Cancer Quality Register. Date of death was merged from the annually updated, high-coverage National Cause of Death Register. Comparisons were made between women in the highest and the lowest quartile of the scores and between the combination of PRS and TCRS, above versus below median, in both scores. Three models were considered: model 1, only adjusting for calendar time; model 2, adjusting for calendar time and age at diagnosis; and model 3, further adjusting for tumor prognosticators. Schoenfeld residuals plotted against time and a χ^2 test were used to assess the proportional hazards assumption.

RESULTS

Descriptive statistics for the cohort are listed in Table 1. Of the 5,500 women, 5,232 had information on TCRS, 4,927 had information on PRS, and 3,488 had a prediagnostic MD measurement. Most women were postmenopausal (70%); the median age was 59 years, with a range of 23 to 79 years. In the cohort, TCRS ranged from 0.23% to 28.86%, with a median of 3.13% (standard deviation [SD], 2.26%). Percent MD ranged from 0.001% to 91.5%, with a median of 19.19% (SD, 13.5%), and PRS ranged from −0.0049 to 0.0154, with a median of 0.005 (SD, 0.002%). Spearman's rank correlation coefficient was −0.004 between TCRS and PRS, −0.12 between TCRS and MD, and 0.04 between PRS and MD.

Table 2 shows the results of binomial and multinomial logistic regression of TCRS, PRS, and MD, overall and stratified by age. TCRS was associated with ER status, grade, and lymph node involvement (ER-negative odds ratio [OR], 0.80; 95% CI, 0.72 to 0.90; lymph node-positive OR, 0.77; 95% CI, 0.68 to 0.87; grade 3 OR, 0.79; 95% CI, 0.69 to 0.90; Table 2), but not with tumor size. Associations remained in women younger than age 50 at diagnosis but were not shown in women aged 50 to 79 years at diagnosis. On finer age stratification, these effects were driven by associations in women aged 40 years or younger (Appendix Table A1, online-only).

The PRS was associated with several prognosticators. Each 1-SD increase in the score was associated with smaller, low-grade, hormone-positive tumors (ER-negative OR, 0.80; 95% CI, 0.74 to 0.87; tumor size greater than 40 mm OR, 0.86; 95% CI, 0.76 to 0.99; grade 3 OR, 0.86; 95% CI, 0.77 to 0.95; Table 2). In analyses stratified by age at diagnosis, patterns of associations were similar in both young and old women. MD was positively associated with lymph node metastasis and tumor size in crude analysis, but showed null associations with grade and ER status. After adjusting

Table 1. Availability and Structure of Data for Age, Tyrer-Cuzick–Predicted 10-Year Risk Score, Polygenic Risk Score, and Mammographic Density in the Cohort

Characteristics	Full Cohort	Women With Data on TCRS	Women With Data on PRS	Women With Data on MD
Total No. of women (% missing data)	5,500 (0)	5,232 (5)	4,927 (10)	3,488 (36)
Age, years				
< 50	1,114 (20)	1,048 (20)	924 (19)	489 (14)
≥ 50	4,386 (80)	4,184 (80)	4,003 (81)	2,999 (86)
No. of postmenopausal women (% of total)	3,778 (70)	3,607 (70)	3,460 (71)	2,603 (75.5)
No. of women with positive family history (% of total)	1,010 (20)	993 (20)	915 (20)	628 (20)
BMI, kg/m^2				
< 20	228 (5)	222 (5)	207 (4)	136 (4)
20-24.9	2,313 (45)	2,279 (45)	2,155 (46)	1,481 (45)
≥ 25	2,580 (50)	2,535 (50)	2,374 (50)	1,647 (51)
HRT use				
Never	2,617 (58)	2,581 (58)	2,389 (57)	1,515 (53)
Past	988 (22)	970 (22)	938 (23)	702 (25)
Current	899 (20)	898 (20)	849 (20)	616 (22)
Estrogen receptor status				
Negative	749 (16)	716 (16)	661 (16)	444 (15)
Positive	3,973 (84)	3,805 (84)	3,568 (84.5)	2,528 (85)
Lymph node status				
Positive	551 (10)	519 (10)	478 (10)	304 (9)
Negative	4,908 (90)	4,675 (90)	4,414 (90)	3,174 (91)
Tumor size, mm				
< 20	3,276 (64)	3,127 (65)	2,960 (65)	2,167 (66)
20-40	1,534 (30)	1,462 (30)	1,361 (30)	935 (29)
> 40	278 (6)	262 (5)	242 (5)	152 (5)
Grade				
1	625 (19)	598 (19)	568 (19)	382 (20)
2	1,694 (52)	1,633 (52)	1,528 (52)	1,034 (53)
3	962 (29)	921 (29)	840 (29)	531 (27)

NOTE. Values are No. (% of total).
Abbreviations: BMI, body mass index; ER, estrogen receptor; HRT, hormone replacement therapy; MD, mammographic density; PRS, polygenic risk score; TCRS, Tyrer-Cuzick-predicted 10-year risk score.

for age at diagnosis, no prognosticator was associated with MD (Table 2). Further adjustments for body mass index, hormone replacement therapy use, and menopausal status did not change estimates (not shown). In analyses stratified by age at diagnosis, no distinct differences were observed (Table 2).

Table 3 shows the results from analyses of TCRS and MD by ER status and of PRS weighted by ER-positive and ER-negative disease. The same pattern was observed for TCRS among ER-positive disease, as in the overall analysis, with a high score being more common in low-grade, lymph node-negative patients. Among ER-negative disease, no prognosticator was associated with TCRS. Analysis using the ER-positive weighted PRS yielded point estimates similar to the main PRS, whereas the ER-negative weighted score was associated with ER-negative receptor status. MD was not associated with any tumor characteristic in any of the strata. There was a nonsignificant trend of high-grade cancer among ER-negative patients with high MD (Table 3).

Sensitivity analysis using 5-year TCRS yielded results similar to the main analysis using 10-year TCRS. The associations for both scores were driven by effects among patients with early-onset disease (Appendix Table A1, online-only). Sensitivity analysis excluding women with a family history of breast cancer revealed more pronounced effect sizes for TCRS, whereas no change was shown for PRS or MD (Appendix Table A2, online-only).

Because TCRS and PRS were not correlated but showed associations between ER and grade in the same direction, we analyzed the combined effect of TCRS and PRS. Women scoring above the median in both TCRS and PRS had more favorable tumor characteristics than did those scoring below the median on both measures (ER-negative OR, 0.52; 95% CI, 0.41 to 0.66; grade 3 OR, 0.56; 95% CI, 0.43 to 0.78; Fig 1).

We further studied the combination of high TCRS and PRS on the risk of being diagnosed with a distant metastasis. For women scoring above the median in both TCRS and PRS, an approximately 30% decreased risk of a distant metastasis was shown (HR, 0.69; 95% CI, 0.49 to 0.98; Table 4). The lowered risk was restricted to patients younger than 50 years of age at diagnosis (HR, 0.20; 95% CI, 0.05 to 0.84), with null associations in women aged 50 or older (Table 4). Neither PRS nor MD were individually associated with a risk of distant metastasis in our cohort, but there was a protective effect for TCRS (HR, 0.69; 95% CI, 0.50 to 0.93), slightly attenuated by adjusting for age and prognosticators (Table 4). No violations of the proportionality assumptions were shown.

DISCUSSION

For breast cancer risk prediction models to aid in lowering disease mortality, they should ideally be able to predict both aggressive and less aggressive subtypes of the disease. We found that women at high risk for breast cancer based on the TCRS or PRS, or both, were significantly more likely to be diagnosed with low-grade, ER-positive cancers. TCRS was also associated with lymph node involvement, and PRS was also influenced by tumor size. The

Table 2. Binomial and Multinomial Logistic Regression of Tyrer-Cuzick–Predicted 10-Year Risk Score, Polygenic Risk Score, and Mammographic Density Versus Tumor Prognosticators: Main Analysis and Stratified by Age

	TCRS			PRS			MD*		
	All Women	Women < 50 years	Women ≥ 50 years	All Women	Women < 50 years	Women ≥ 50 years	All Women	Women < 50 years	Women ≥ 50 years
ER status									
Positive	1.00 (ref)	1.00 (ref)	1.00 (ref)	1.00 (ref)	1.00 (ref)	1.00 (ref)	1.00 (ref)	1.00 (ref)	1.00 (ref)
Negative	**0.80 (0.72 to 0.90)**	**0.57 (0.39 to 0.81)**	0.91 (0.81 to 1.02)	**0.80 (0.74 to 0.87)**	**0.78 (0.66 to 0.93)**	**0.80 (0.73 to 0.88)**	0.93 (0.83 to 1.04)	0.81 (0.62 to 1.07)	0.95 (0.81 to 1.07)
Lymph node status									
Negative	1.00 (ref)	1.00 (ref)	1.00 (ref)	1.00 (ref)	1.00 (ref)	1.00 (ref)	1.00 (ref)	1.00 (ref)	1.00 (ref)
Positive	**0.77 (0.68 to 0.87)**	0.76 (0.55 to 1.01)	0.94 (0.82 to 1.06)	0.98 (0.91 to 1.10)	0.87 (0.73 to 1.04)	1.03 (0.92 to 1.15)	1.11 (0.98 to 1.25)	1.02 (0.80 to 1.30)	1.16 (0.99 to 1.33)
Tumor size									
< 20 mm	1.00 (ref)	1.00 (ref)	1.00 (ref)	1.00 (ref)	1.00 (ref)	1.00 (ref)	1.00 (ref)	1.00 (ref)	1.00 (ref)
20-40 mm	0.96 (0.90 to 1.03)	0.81 (0.64 to 1.03)	1.02 (0.95 to 1.10)	0.94 (0.88 to 1.00)	1.05 (0.90 to 1.21)	**0.91 (0.85 to 0.98)**	1.03 (0.95 to 1.12)	**0.75 (0.61 to 0.92)**	1.11 (1.02 to 1.21)
> 40 mm	0.97 (0.84 to 1.11)	**0.50 (0.29 to 0.87)**	1.13 (0.99 to 1.28)	**0.86 (0.76 to 0.99)**	**0.77 (0.58 to 0.98)**	0.90 (0.77 to 1.05)	1.03 (0.87 to 1.23)	1.15 (0.79 to 1.69)	0.99 (0.81 to 1.20)
Grade									
1	1.00 (ref)	1.00 (ref)	1.00 (ref)	1.00 (ref)	1.00 (ref)	1.00 (ref)	1.00 (ref)	1.00 (ref)	1.00 (ref)
2	1.03 (0.92 to 1.15)	0.96 (0.62 to 1.47)	1.11 (0.98 to 1.26)	1.03 (0.94 to 1.13)	1.12 (0.86 to 1.46)	1.00 (0.91 to 1.11)	1.02 (0.87 to 1.16)	0.88 (0.59 to 1.32)	1.04 (0.91 to 1.19)
3	**0.79 (0.69 to 0.90)**	0.73 (0.46 to 1.15)	0.94 (0.81 to 1.10)	**0.86 (0.77 to 0.95)**	**0.77 (0.59 to 1.00)**	**0.87 (0.77 to 0.97)**	0.87 (0.75 to 1.01)	0.82 (0.53 to 1.25)	0.86 (0.74 to 1.01)

NOTE. Values are OR (95% CI). ORs with 95% CIs are shown per 1-SD increase. Boldface type indicates associations significant at alpha = 0.05.
Abbreviations: ER, estrogen receptor; MD, mammographic density; OR, odds ratio; PRS, polygenic risk score; ref, reference; SD, standard deviation; TCRS, Tyrer-Cuzick-predicted 10-year risk score.
*Age-adjusted.

Table 3. Binomial and Multinomial Logistic Regression of Tyrer-Cuzick–Predicted 10-Year Risk Score and Mammographic Density Stratified by ER Status and Polygenic Risk Score Weighted on ER-Positive Disease and ER-Negative Disease

	TCRS			PRS*			MD†		
	All Women	ER-Positive	ER-Negative	Overall	ER-Positive	ER-Negative	All Women	ER-Positive	ER-Negative
ER status									
Positive	1.00 (ref)			1.00 (ref)	1.00 (ref)	1.00 (ref)	1.00 (ref)		
Negative	**0.80 (0.72 to 0.90)**			**0.80 (0.74 to 0.87)**	**0.76 (0.70 to 0.83)**	**1.14 (1.05 to 1.24)**	0.93 (0.83 to 1.04)		
Lymph node status									
Negative	1.00 (ref)	1.00 (ref)	1.00 (ref)	1.00 (ref)	1.00 (ref)	1.00 (ref)	1.00 (ref)	1.00 (ref)	1.00 (ref)
Positive	**0.77 (0.68 to 0.87)**	**0.76 (0.65 to 0.89)**	1.01 (0.78 to 1.26)	0.98 (0.91 to 1.10)	0.99 (0.90 to 1.09)	1.03 (0.93 to 1.13)	1.11 (0.98 to 1.25)	1.13 (0.97 to 1.30)	1.28 (0.97 to 1.69)
Tumor size									
< 20 mm	1.00 (ref)	1.00 (ref)	1.00 (ref)	1.00 (ref)	1.00 (ref)	1.00 (ref)	1.00 (ref)	1.00 (ref)	1.00 (ref)
20-40 mm	0.96 (0.90 to 1.03)	0.96 (0.88 to 1.05)	0.92 (0.77 to 1.11)	0.94 (0.88 to 1.00)	**0.92 (0.87 to 0.98)**	1.07 (0.99 to 1.14)	1.03 (0.95 to 1.12)	1.07 (0.97 to 1.18)	1.02 (0.83 to 1.27)
> 40 mm	0.97 (0.84 to 1.11)	1.05 (0.88 to 1.04)	0.97 (0.69 to 1.36)	**0.86 (0.76 to 0.99)**	**0.86 (0.75 to 0.98)**	0.96 (0.84 to 1.09)	1.03 (0.87 to 1.23)	1.11 (0.90 to 1.38)	0.96 (0.64 to 1.46)
Grade									
1	1.00 (ref)	1.00 (ref)	1.00 (ref)	1.00 (ref)	1.00 (ref)	1.00 (ref)	1.00 (ref)	1.00 (ref)	1.00 (ref)
2	1.03 (0.92 to 1.15)	1.03 (0.91 to 1.15)	1.38 (0.59 to 3.02)	1.03 (0.94 to 1.13)	1.02 (0.93 to 1.13)	1.05 (0.95 to 1.15)	1.02 (0.87 to 1.16)	1.03 (0.90 to 1.18)	1.85 (0.60 to 5.75)
3	**0.79 (0.69 to 0.90)**	**0.84 (0.72 to 0.98)**	1.03 (0.45 to 2.35)	**0.86 (0.77 to 0.95)**	**0.82 (0.74 to 0.92)**	1.11 (1.00 to 1.24)	0.87 (0.75 to 1.01)	0.84 (0.74 to 1.00)	2.10 (0.70 to 6.36)

NOTE. Values are OR (95% CI). Odds ratios with 95% CIs are shown per 1-SD increase. Boldface type indicates associations significant at alpha = 0.05.
Abbreviations: ER, estrogen receptor; MD, mammographic density; OR, odds ratio; PRS, polygenic risk score; ref, reference; SD, standard deviation; TCRS, Tyrer-Cuzick-predicted 10-year risk score.
*All analysis was conducted for all patients.
†Age-adjusted.

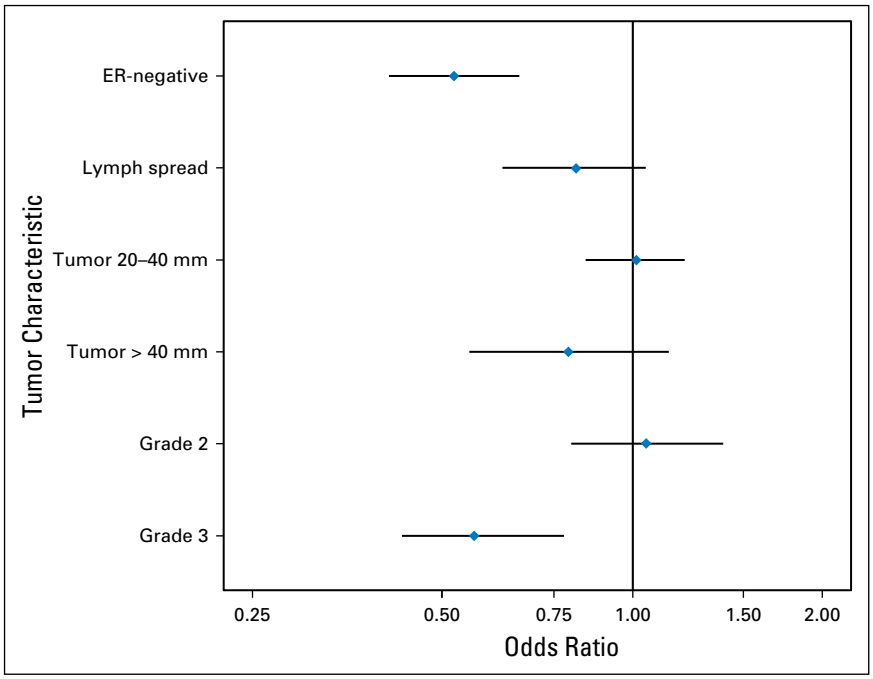

Fig 1. Associations with tumor prognosticators in women scoring above the median in polygenic risk score (PRS) and Tyrer-Cuzick–predicted 10-year risk score (TCRS) versus women scoring below the median in PRS and TCRS. Shown are odds ratios with 95% confidence intervals. ER, estrogen receptor.

influence of PRS was age independent, in contrast to TCRS, which was only associated with tumor characteristics in young women. We further found that the risk of being diagnosed with distant metastasis was significantly lower in women at increased risk for breast cancer, according to both TCRS and PRS. MD showed null associations with all tumor prognosticators under study and seems therefore to be a general risk factor for breast cancer.

To the best of our knowledge, no study has previously addressed how tumor prognosticators vary with the predicted breast cancer risk using the Tyrer-Cuzick model. Several of the lifestyle and reproductive risk factors determining the score have been shown to be positively associated with ER-positive disease,[10] in agreement with our results. We found that patients with a high TCRS were more likely to have ER-positive tumors, and the largest effect sizes were shown when excluding patients with a family history of breast cancer, thus restricting the score to signifying hormonal and lifestyle risk factors. Stratified analysis revealed that these main effects were driven by associations in women younger than 40 years of age. This suggests that risk factors incorporated in the Tyrer-Cuzick model are less useful when aiming to identify aggressive breast cancer in young women, but may work well as a general tool for postmenopausal cancers. Moreover, if high-risk young women according to TCRS are at risk for developing ER-positive subtypes, TCRS could be a tool for identifying healthy

Table 4. Cox Regression Modeling of Hazard Ratios of Distant Metastasis

	No. of Events	All Women	No. of Events	Women < 50 years	No. of Events	Women ≥ 50 years
Model 1*						
TCRS	164	**0.69 (0.50 to 0.93)**	69	0.83 (0.38 to 1.82)	95	0.83 (0.54 to 1.25)
PRS	145	0.91 (0.66 to 1.26)	38	0.66 (0.35 to 1.25)	107	0.99 (0.68 to 1.45)
MD	85	1.13 (0.74 to 1.74)	18	0.50 (0.14 to 1.73)	67	1.17 (0.72 to 1.90)
TCRS/PRS combined	129	**0.69 (0.49 to 0.98)**	39	**0.20 (0.05 to 0.84)**	90	1.02 (0.66 to 1.53)
Model 2†						
TCRS	164	0.74 (0.53 to 1.06)	69	0.98 (0.44 to 2.20)	95	0.82 (0.54 to 1.25)
PRS	145	0.90 (0.65 to 1.25)	38	0.66 (0.35 to 1.25)	107	0.99 (0.68 to 1.45)
MD	85	1.10 (0.68 to 1.76)	18	0.47 (0.13 to 1.64)	67	1.22 (0.74 to 2.02)
TCRS/PRS combined	129	0.80 (0.56 to 1.15)	39	**0.23 (0.06 to 0.98)**	90	1.01 (0.66 to 1.53)
Model 3‡						
TCRS	164	0.73 (0.51 to 1.04)	69	1.10 (0.48 to 2.49)	95	0.68 (0.44 to 1.04)
PRS	145	0.91 (0.65 to 1.26)	38	0.66 (0.35 to 1.26)	107	1.04 (0.71 to 1.52)
MD	85	1.04 (0.64 to 1.69)	18	0.44 (0.12 to 1.65)	67	1.15 (0.69 to 1.51)
TCRS/PRS combined	129	0.81 (0.56 to 1.18)	39	0.26 (0.06 to 1.10)	90	0.93 (0.61 to 1.43)

NOTE. Values are No. of events or HR (95% CI). HRs are shown for overall analysis and stratified to old and young women. TCRS, PRS, and MD HRs represent women in the top quartile of the scores being compared with women in the lowest quartiles. For TCRS/PRS combined, HRs represent women above median of TCRS and PRS are compared with women below median of TCRS and PRS. Boldface type indicates associations significant at alpha = 0.05.

Abbreviations: HR, hazard ratio; MD, mammographic density; PRS, polygenic risk score; TCRS, Tyrer-Cuzick–predicted 10-year risk score.

*Adjusted for calendar period.

†Adjusted for calendar period and age (continuous).

‡Adjusted for calendar period, age (continuous), estrogen receptor status, lymph node involvement, and tumor size. For TCRS, also adjusted for grade.

women who could be offered prophylactic treatment with anti-estrogens, such as tamoxifen.

Interestingly, associations with grade and lymph node status remained within the strata of ER-positive cancers; thus, we observed an effect beyond ER status. High-grade ER-positive cancers are more often of the luminal B subtype, a type found with younger age of onset, whereas low-grade ER-positive tumors are commonly of the subtype luminal A.[18] Given the over-representation of luminal A breast cancers in most populations in which the etiology has been studied, a bias among currently known risk factors toward luminal A might explain the associations. The distribution of TCRS within molecular subtypes is of great interest for addressing this hypothesis in future work.

The associations between the 77-SNP PRS and tumor prognosticators are in agreement with the one previous study addressing a similar question.[19] In contrast to TCRS, the PRS findings were independent of age. This would be expected with genetic markers if they predict disease independent of age-of-onset patterns. Beyond ER status, we also found an association with tumor grade. This is not altogether surprising, because there is a modest correlation between ER status and grade. When stratifying our analysis by ER status, all associations with grade disappeared within both strata (data not shown). The reason for the PRS bias toward ER-positive tumors is likely because of the abundance of ER-positive cancers in SNP discovery cohorts. Studies designed to challenge this issue have identified SNPs associated with ER-negative disease,[20] and the ER-negative weighted PRS would be an important tool to examine, in addition to the overall PRS, when assessing individual risk. We found ER-negative weighted PRS to be associated with ER-negative disease, thus showing potential as a complement to the overall PRS score for identification of healthy women at risk for aggressive disease.

Of the three risk prediction tools characterized in this study, MD has been studied most thoroughly for associations with tumor prognosticators. Our results are consistent with the outcome of a meta-analysis on MD and ER status, showing null associations,[21] and the null associations with lymph node involvement and size, after adjusting for age, are consistent with most other published literature.[22] In our study, we could not replicate the previous finding of an association between MD and an increased risk of ER-negative breast cancer in young women.[23]

Both PRS and TCRS were associated with favorable tumor prognosticators, and women scoring high in both tools may be even more likely to have favorable disease outcomes, as judged by effect sizes from combinatorial analysis. The survival analysis supported these findings because women above the median in both PRS and TCRS were at decreased risk of distant metastasis. Again, the effects were constrained to women younger than age 50. In this group, adjusting for standard prognosticators did not alter the effect size substantially, although the statistical significance was lost. Given that young patients are at highest risk of metastasis,[24] our results warrant further investigation in future studies. However, this is the first study of TCRS/PRS and survival, and this part of our results should be interpreted with caution until replication in independent cohorts can be performed. Our survival analysis suggests that young women at risk for aggressive disease may be classified as low-risk women by both TCRS/PRS prediction tools.

Our study has limitations that should be discussed. The survivor bias arising from including women alive in 2009 might influence our findings. We performed sensitivity analysis, restricting to women diagnosed between 2004 and 2008, but saw no considerable changes to point estimates. Moreover, we had no information of history of benign breast disease, one of the components of the Tyrer-Cuzick model. Although we do not believe that such a history is associated with any of the tumor characteristics under study, this should be kept in mind when comparing with future studies. Family history was restricted to first-degree relatives, which could underestimate the effects of family history on risk. However, because the Tyrer-Cuzick shows a high-penetrance genetic risk, the model should not be sensitive to the loss of information on distant relatives. Finally, our results are representative of a homogenous Scandinavian population and may not translate to other ethnic groups and geographical regions.

The main strengths of this work lie in the high quality of the Swedish health registers, combined with the availability of automated MD estimates, questionnaire data, and genotype information for each patient. We addressed our study question in a population-based cohort of breast cancer patients with a large number of participants, thereby generating high external validity. The regional cancer register provides near-complete information on tumor prognosticators and good follow-up for distant metastasis; moreover, the completeness of the cause of death register ensures reliable censoring information, and the reader-independent MD measurements limit any observer bias.

In conclusion, to our knowledge, we showed for the first time that a high risk of breast cancer, based on TCRS and PRS, is associated with favorable tumor prognosticators and a reduced risk of being diagnosed with distant metastasis. PRS weighted on ER-negative disease should be considered for identifying women at risk for aggressive disease. Our results support the hypothesis that breast cancer subtypes have different etiologies and highlight the need to identify risk factors separately for distinct breast cancer subtypes and ages of onset. Better knowledge of subtype-specific risk factors and understanding of disease etiology may be vital for the success of primary prevention and screening programs aimed at lowering mortality.

AUTHORS' DISCLOSURES OF POTENTIAL CONFLICTS OF INTEREST

Disclosures provided by the authors are available with this article at www.jco.org

AUTHOR CONTRIBUTIONS

Conception and design: Johanna Holm, Keith Humphreys, Per Hall, Kamila Czene
Financial support: Kamila Czene
Administrative support: Kamila Czene
Provision of study materials or patients: Kamila Czene
Collection and assembly of data: Johanna Holm, Jingmei Li, Mikael Eriksson, Per Hall, Kamila Czene
Data analysis and interpretation: Johanna Holm, Jingmei Li, Hatef Darabi, Martin Eklund, Keith Humphreys, Per Hall, Kamila Czene
Manuscript writing: All authors
Final approval of manuscript: All authors

REFERENCES

1. Howell A, Anderson AS, Clarke RB, et al: Risk determination and prevention of breast cancer. Breast Cancer Res 16:446, 2014
2. Tyrer J, Duffy SW, Cuzick J: A breast cancer prediction model incorporating familial and personal risk factors. Stat Med 23:1111-1130, 2004
3. Mavaddat N, Pharoah PDP, Michailidou K, et al: Prediction of breast cancer risk based on profiling with common genetic variants. J Natl Cancer Inst 107:djv036, 2015
4. McCormack VA, dos Santos Silva I: Breast density and parenchymal patterns as markers of breast cancer risk: a meta-analysis. Cancer Epidemiol Biomarkers Prev 15:1159-1169, 2006
5. Warwick J, Birke H, Stone J, et al: Mammographic breast density refines Tyrer-Cuzick estimates of breast cancer risk in high-risk women: findings from the placebo arm of the International Breast Cancer Intervention Study I. Breast Cancer Res 16:451, 2014
6. Darabi H, Czene K, Zhao W, et al: Breast cancer risk prediction and individualised screening based on common genetic variation and breast density measurement. Breast Cancer Res 14:R25, 2012
7. Sestak I, Cuzick J: Update on breast cancer risk prediction and prevention. Curr Opin Obstet Gynecol 27:92-97, 2015
8. Vachon CM, Pankratz VS, Scott CG, et al: The contributions of breast density and common genetic variation to breast cancer risk. J Natl Cancer Inst 107: dju397, 2015
9. Kwan ML, Kushi LH, Weltzien E, et al: Epidemiology of breast cancer subtypes in two prospective cohort studies of breast cancer survivors. Breast Cancer Res 11:R31, 2009
10. Yang XR, Chang-Claude J, Goode EL, et al: Associations of breast cancer risk factors with tumor subtypes: a pooled analysis from the Breast Cancer Association Consortium studies. J Natl Cancer Inst 103:250-263, 2011
11. Millikan RC, Newman B, Tse C-K, et al: Epidemiology of basal-like breast cancer. Breast Cancer Res Treat 109:123-139, 2008
12. Emilsson L, Lindahl B, Köster M, et al: Review of 103 Swedish Healthcare Quality Registries. J Intern Med 277:94-136, 2015
13. Michailidou K, Hall P, Gonzalez-Neira A, et al: Large-scale genotyping identifies 41 new loci associated with breast cancer risk. Nat Genet 45:353-361, 361e1-361e2, 2013
14. Li J, Holm J, Bergh J, et al: Breast cancer genetic risk profile is differentially associated with interval and screen-detected breast cancers. Ann Oncol 26:517-522, 2015
15. Li J, Szekely L, Eriksson L, et al: High-throughput mammographic-density measurement: a tool for risk prediction of breast cancer. Breast Cancer Res 14:R114, 2012
16. Byng JW, Boyd NF, Fishell E, et al: The quantitative analysis of mammographic densities. Phys Med Biol 39:1629-1638, 1994
17. R Development Core Team R: R: A language and environment for statistical computing. R Found Stat Comput 1:409, 2011
18. Prat A, Cheang MCU, Martín M, et al: Prognostic significance of progesterone receptor-positive tumor cells within immunohistochemically defined luminal A breast cancer. J Clin Oncol 31:203-209, 2013
19. Reeves GK, Travis RC, Green J, et al: Million Women Study Collaborators: Incidence of breast cancer and its subtypes in relation to individual and multiple low-penetrance genetic susceptibility loci. JAMA 304:426-434, 2010
20. Purrington KS, Slager S, Eccles D, et al: GENICA Network: Genome-wide association study identifies 25 known breast cancer susceptibility loci as risk factors for triple-negative breast cancer. Carcinogenesis 35:1012-1019, 2014
21. Antoni S, Sasco AJ, dos Santos Silva I, et al: Is mammographic density differentially associated with breast cancer according to receptor status? A meta-analysis. Breast Cancer Res Treat 137:337-347, 2013
22. Boyd NF, Martin LJ, Yaffe MJ, et al: Mammographic density and breast cancer risk: Current understanding and future prospects. Breast Cancer Res 13:223, 2011
23. Bertrand KA, Tamimi RM, Scott CG, et al: Mammographic density and risk of breast cancer by age and tumor characteristics. Breast Cancer Res 15: R104, 2013
24. Purushotham A, Shamil E, Cariati M, et al: Age at diagnosis and distant metastasis in breast cancer–a surprising inverse relationship. Eur J Cancer 50: 1697-1705, 2014

GLOSSARY TERMS

Cox proportional hazards regression model: a statistical model for regression analysis of censored survival data, examining the relationship of censored survival distribution to one or more covariates. This model produces a baseline survival curve, covariate coefficient estimates with their standard errors, risk ratios, 95% CIs, and significance levels.

hazard ratios: the ratio of the hazard rate in one group (eg, a group of treated patients) to the hazard rate in another group (eg, an untreated control group of patients). The hazard rate is the probability of a specified event, such as death or cancer recurrence, occurring during a short time interval. The hazard ratio, therefore, is a measure of the relative probability of an event occurring at any given point in time.

logistic regression analysis: a multivariable regression model in which the log of the odds of a time-fixed outcome event (eg, 30-day mortality) or other binary outcome is related to a linear equation.

AUTHORS' DISCLOSURES OF POTENTIAL CONFLICTS OF INTEREST

Association of Breast Cancer Risk Prediction Tools With Tumor Characteristics and Metastasis

The following represents disclosure information provided by authors of this manuscript. All relationships are considered compensated. Relationships are self-held unless noted. I = Immediate Family Member, Inst = My Institution. Relationships may not relate to the subject matter of this manuscript. For more information about ASCO's conflict of interest policy, please refer to www.asco.org/rwc or jco.ascopubs.org/site/ifc.

Johanna Holm
No relationship to disclose

Jingmei Li
No relationship to disclose

Hatef Darabi
No relationship to disclose

Martin Eklund
Employment: Genetta Soft, Metasafe AB (I)
Stock or Other Ownership: Genetta Soft, Metasafe AB (I)
Patents, Royalties, Other Intellectual Property: Procant

Mikael Eriksson
No relationship to disclose

Keith Humphreys
No relationship to disclose

Per Hall
No relationship to disclose

Kamila Czene
No relationship to disclose

Acknowledgment

We thank Louise Eriksson, PhD, for her valuable input and careful reading of the manuscript.

Appendix

Table A1. Binomial and Multinomial Logistic Regression of TCRS and TCRS5 Versus Tumor Prognosticators, Stratified by Age: ≤ 40, 41 to 50, 51 to 60, > 60 Modeled per 1-SD Increase

	TCRS				TCRS5			
	Women ≤ 40 (n = 298)	Women 41 to 50 (n = 1,063)	Women 51 to 60 (n = 1,785)	Women > 60 (n = 2,354)	Women ≤ 40 (n = 298)	Women 41 to 50 (n = 1,063)	Women 51 to 60 (n = 1,785)	Women > 60 (n = 2,354)
ER positive	1.00 (Ref)	1.00 (Ref)	1.00 (Ref)	1.00 (Ref)	1.00 (Ref)	1.00 (Ref)	1.00 (Ref)	1.00 (Ref)
ER negative	**0.37 (0.15 to 0.86)**	0.99 (0.67 to 1.43)	0.90 (0.75 to 1.07)	0.89 (0.74 to 1.05)	**0.20 (0.05 to 0.65)**	0.93 (0.62 to 1.44)	0.90 (0.74 to 1.07)	0.89 (0.74 to 1.05)
Lymph node negative	1.00 (Ref)	1.00 (Ref)	1.00 (Ref)	1.00 (Ref)	1.00 (Ref)	1.00 (Ref)	1.00 (Ref)	1.00 (Ref)
Lymph node positive	**0.31 (0.10 to 0.81)**	0.86 (0.62 to 1.11)	0.95 (0.77 to 1.12)	0.94 (0.78 to 1.10)	**0.19 (0.04 to 0.70)**	0.85 (0.59 to 1.13)	0.95 (0.76 to 1.13)	0.94 (0.78 to 1.09)
Tumor < 20 mm	1.00 (Ref)	1.00 (Ref)	1.00 (Ref)	1.00 (Ref)	1.00 (Ref)	1.00 (Ref)	1.00 (Ref)	1.00 (Ref)
Tumor 20 to 40 mm	0.90 (0.42 to 1.91)	0.89 (0.73 to 1.10)	1.08 (0.97 to 1.21)	0.96 (0.86 to 1.07)	0.94 (0.34 to 2.62)	0.89 (0.71 to 1.11)	1.08 (0.97 to 1.22)	0.95 (0.85 to 1.06)
Tumor > 40 mm	1.04 (0.33 to 3.27)	0.69 (0.39 to 1.23)	1.05 (0.84 to 1.32)	1.14 (0.97 to 1.33)	1.20 (0.26 to 5.52)	0.67 (0.35 to 1.26)	1.06 (0.83 to 1.35)	1.16 (0.99 to 1.36)
Grade 1	1.00 (Ref)	1.00 (Ref)	1.00 (Ref)	1.00 (Ref)	1.00 (Ref)	1.00 (Ref)	1.00 (Ref)	1.00 (Ref)
Grade 2	0.61 (0.13 to 2.87)	1.18 (0.72 to 1.92)	**1.29 (1.03 to 1.63)**	1.10 (0.87 to 1.18)	0.45 (0.06 to 3.45)	1.21 (0.70 to 2.08)	**1.30 (1.02 to 1.65)**	1.00 (0.86 to 1.16)
Grade 3	0.70 (0.16 to 3.00)	1.35 (0.81 to 2.25)	1.00 (0.77 to 1.30)	0.89 (0.73 to 1.07)	0.45 (0.07 to 3.14)	1.37 (0.78 to 2.43)	0.95 (0.72 to 1.26)	0.86 (0.71 to 1.03)

NOTE. Values are OR (95% CI). Boldface type indicates associations significant at alpha = 0.05.
Abbreviations: ER, estrogen receptor; OR, odds ratio; Ref, reference; SD, standard deviation; TCRS, Tyrer-Cuzick-predicted 10-year risk; TCRS5, Tyrer-Cuzick-predicted 5-year risk.

Table A2. Binomial and Multinomial Logistic Regression of Each Tool Versus Tumor Prognosticators in Women Without Family History: TCRS, PRS, and MD Modeled per 1-SD Increase

	TCRS	PRS	MD
ER positive	1.00 (Ref)	1.00 (Ref)	1.00 (Ref)
ER negative	**0.71 (0.56 to 0.89)**	**0.77 (0.70 to 0.84)**	0.92 (0.81 to 1.04)
Lymph node negative	1.00 (Ref)	1.00 (Ref)	1.00 (Ref)
Lymph node positive	**0.43 (0.33 to 0.56)**	1.01 (0.91 to 1.13)	1.07 (0.93 to 1.24)
Tumor < 20 mm	1.00 (Ref)	1.00 (Ref)	1.00 (Ref)
Tumor 20 to 40 mm	0.96 (0.87 to 1.07)	0.95 (0.88 to 1.03)	1.07 (0.97 to 1.18)
Tumor > 40 mm	**0.70 (0.48 to 0.99)**	**0.82 (0.70 to 0.95)**	0.95 (0.77 to 1.18)
Grade 1	1.00 (Ref)	1.00 (Ref)	1.00 (Ref)
Grade 2	1.01 (0.81 to 1.27)	1.03 (0.93 to 1.15)	0.98 (0.85 to 1.13)
Grade 3	**0.43 (0.33 to 0.59)**	**0.85 (0.76 to 0.96)**	0.89 (0.75 to 1.05)

NOTE. Values are OR (95% CI). MD is adjusted for age. Boldface type indicates associations significant at alpha = 0.05.
Abbreviations: ER, estrogen receptor; MD, mammographic density; OR, odds ratio; PRS, polygenic risk score; Ref, reference; SD, standard deviation; TCRS, Tyrer-Cuzick predicted 10-year risk.