The path to individualised breast cancer screening

Lakeman, I.M.M.

CHAPTER 3

3

Clinical applicability of the Polygenic Risk Score for breast cancer risk prediction in familial cases

Inge M.M. Lakeman, Mar Rodriguez-Girondo, Andrew Lee, Nandi Celosse, Merel Braspenning, K. van Engelen, I. van de Beek, A.H. van der Hout, E. Gomez Garcia, A. Mensenkamp, M.G.E.M. Ausems, Maartje Hooning, Muriel A. Adank, Antoinette Hollestelle, Marjanka K. Schmidt, Christi J. van Asperen, and Peter Devilee

Lakeman IMM, et al. Submitted

Abstract

Background: Common low-risk variants are presently not used to guide clinical management of familial breast cancer (BC). We explored the additive impact of a 313-variant-based Polygenic Risk Score (PRS_{313}) relative to standard gene-testing in non-*BRCA1/2* Dutch BC families.

Methods: We included 3,918 BC cases from 3,492 Dutch non-*BRCA1/2* BC families and 3,474 Dutch population controls. The association of the standardised PRS_{313} with BC was estimated using a logistic regression model, adjusted for pedigree-based family history. Family history of controls was imputed for this analysis. Standard errors were corrected to account for relatedness of individuals. Using BOADICEA model version 5, lifetime risks were retrospectively calculated with and without individual PRS_{313}. For 2,586 cases and 2,584 controls, carrier status of pathogenic variants (PVs) in *ATM*, *CHEK2,* and *PALB2* was known.

Results: The family history adjusted PRS_{313} was significantly associated with BC (per SD OR=1.97, 95%CI[1.84-2.11]). Including the PRS_{313} in BOADICEA family-based risk prediction would have changed screening recommendations in up to 27%, 36%, and 34% of the cases according to BC screening guidelines from the USA, UK and the Netherlands (NCCN, NICE, and IKNL), respectively. For the population controls, without information on family history, this was up to 39%, 44%, and 58%, respectively. Among carriers of PVs in known moderate BC susceptibility genes, the PRS_{313} had the largest impact for *CHEK2* and *ATM*.

Conclusions: Our results support the application of the PRS_{313} in risk prediction for genetically uninformative BC families and families with a PV in moderate BC risk genes.

Introduction

Breast cancer (BC) is the most common cancer among women[1]. Current screening strategies to reduce the burden of the disease have several disadvantages, including overdiagnosis[2]. By taking into account all relevant risk factors, personalised estimation of BC risk could help to target preventive measures to those who would benefit the most and to reduce screening for women in the lowest risk categories.

One of the main BC risk factors is having a positive family history of the disease[3]. The familial relative risk of ~2 is partly explained by germline pathogenic variants (PVs) in the BC susceptibility genes *BRCA1/2, PALB2, ATM* and *CHEK2*. Furthermore, another important part is explained by common low-risk variants[4, 5], which, if summarised in a Polygenic Risk Score (PRS), are useful for stratifying the population into different risk categories[5, 6]. A similar stratification of BC risk by the PRS is observed in the familial setting[7-10], providing an opportunity to personalising risk and clinical management for women from BC families who are seen at clinical genetic services. Furthermore, the PRS can be useful in refining risk for women carrying a PV in *BRCA1/2, PALB2, CHEK2,* or *ATM*[11-14]. However, using the PRS for risk prediction is not yet implemented in the practice of genetic counselling for familial BC in the Netherlands.

Currently, risk prediction for women from non-*BRCA1/2* BC families is mainly based on family history, which can be calculated by various risk prediction algorithms[15], such as the Breast and Ovarian Analysis of Disease Incidence and Carrier Estimation Algorithm (BOADICEA)[16]. Several studies have shown an improved discriminative power between BC cases and controls by combining the PRS with other risk factors in a BC risk prediction tool[17-20]. Previously, we showed that in a selected group of high risk non-*BRCA1/2* BC families, a 161-variant PRS alone would have led 20% of the women to receive different screening recommendations based on the Dutch screening guideline (Netherlands Comprehensive Cancer Organisation guideline (IKNL))[21]. Currently, the most predictive PRS, based on 313 variants (PRS_{313})[5], is incorporated in the validated, comprehensive risk prediction model BOADICEA[16] that was recently made easily accessible for clinicians through the CanRisk webtool[22].

Here, we explore the clinical applicability of the PRS_{313} for risk prediction in a new cohort of 3,918 familial Dutch BC cases who tested negative in a diagnostic setting for PVs in *BRCA1/2* and of whom the majority were evaluated for PVs in *PALB2, CHEK2*, and *ATM* in a research setting. The clinical impact of the PRS_{313} on BC risk prediction based on family history and PV carrier status was investigated by determining the potential change in clinical management, as stipulated by three currently used guidelines (the National

Comprehensive Cancer Network guideline (NCCN)[23], the National Institute for Health and Care Excellence guideline (NICE)[24], and IKNL[21]).

Materials and Methods

We used the STROBE case-control checklist when writing our report[25].

Study cohorts

Dutch familial BC cases, henceforth "cases", were derived from three different cohorts: the Hereditary Breast and Ovarian cancer study in the Netherlands (HEBON)[26], the Amsterdam Breast Cancer Study-Familial (ABCS-F)[27], and the Rotterdam Breast Cancer Study (RBCS)[28] (Supplementary methods). All three studies included participants who visited a clinical genetic centre in the Netherlands for familial BC counselling. Women with BC who met the following criteria were eligible for this study: 1) family without *BRCA1/2* PVs; 2) available DNA sample or genotyping data; 3) European ancestry based on genotyping data; 4) available pedigree. In total, 3,918 cases were included (Figure S1). All cancers were verified by linkage to the Dutch Cancer Registry and the Pathological Anatomical National Automated Archive (HEBON cases) or by clinical confirmation from medical records in the hospital (ABCS-F and RBCS cases).

In total, 3,474 Dutch population controls of age 18 years or older were included. These controls were healthy female blood donors (ABCS, Oorsprong van borstkanker integraal onderzocht (ORIGO)) or healthy women who were included after DNA diagnostic testing for Cystic Fibrosis carrier status (RBCS)[4, 28] for which age of last follow up was known.

Ethics approval statement

Informed consent was obtained from all included cases, and we received approval for this study of the Medical Ethical Committees of all included centres. All controls were anonymised.

Gene panel

As part of the BRIDGES project, 2,586 cases and 2,584 controls were sequenced for a panel of 34 genes as described elsewhere[29]. For all controls and 2,037 cases, we received results of all included genes. Truncating and missense variants were reported as described previously[29]. In summary, pathogenic truncating variants were defined as frameshift insertions/deletions, stop/gain or canonical splice variants as classified by the Ensembl Variant Effect Predictor[30], with the exception of variants in the last exon of each gene. In our study, we included truncating variants in the last exon of *PALB2*, as this exon encodes an important functional domain and variants in this exon were shown to destabilise

the resulting *PALB2* protein[31]. Missense variants were included if their frequency in the gnomAD database or among the BRIDGES project control dataset[29] was below 0.001. For genes with evidence of an association with BC[29], pathogenicity was reported for missense variants based on the ClinVar archive[32]. For the remaining 549 cases, only pseudo-anonymised results of truncating variants in the three additional BC genes, *ATM, CHEK2*, and *PALB2,* were received, excluding truncating variants in the last exon.

Genotyping and imputation

DNA samples of all included individuals were genotyped for common variants with either the iCOGS[33], OncoArray[4] or Global Screening Array (GSA), containing 211,155, 499,170, and 642,824 Single Nucleotide Polymorphisms, respectively. Genotyping and quality control for the samples genotyped with iCOGS and OncoArray were performed as part of association studies conducted by the Breast Cancer Association Consortium (BCAC)[4, 33]. Genotyping and quality control for the samples genotyped with the GSA array are described in the supplementary methods.

The variants that were not directly genotyped were imputed using the Michigan imputation server[34], using the Haplotype Reference Consortium (HRC) 1.1 reference panel[35] including both the reference panels 1000 Genomes phase 3 and Genome of the Netherlands (GoNL)[36, 37]. In total, 72 of the 313 variants could not be imputed with the HRC1.1 reference panel and were imputed with the 1000 Genomes phase 3 reference panel only[37] (Table S1).

Polygenic Risk Score

The PRS was calculated as described previously[5]. The three PRSs (for overall BC, ER-positive, and ER-negative BC) were calculated for all included individuals. The variants and their corresponding weights used in the PRS as published previously[5] and the imputation quality are listed in Table S1. The PRS for each individual was standardised to the mean from all population controls in this study and to the SD in the Breast Cancer Association Consortium (BCAC) population controls that were included in the validation data set[5]. These SDs were 0.6093, 0.6520, and 0.5920 for the overall BC PRS, ER-positive BC PRS, and ER-negative BC PRS, respectively. Using these SDs, the OR estimates for the associations of the standardised PRS_{313} in our study are directly comparable with the OR estimates reported in the BCAC population-based study[5].

Pedigree collection

Pedigrees were collected for all families and were drawn previously in the clinical genetic centres during counselling and DNA diagnostic testing of *BRCA1/2* PVs. The pedigrees were used as they were drawn in the clinic, including at least all known first- and second-

degree relatives of the genotyped individuals. Imputation of missing data is described in the supplementary material.

Family history score

A model-based family history score for BC, also called the 'polygenic load', was derived from the BOADICEA version 3 model based on the available pedigree, as described previously[7]. The polygenic load in BOADICEA is a latent polygenetic component representing the combined effect of a large number of variants each of small effect to capture the residual familial aggregation of BC and is, therefore, a measure of the BC family history[7, 10]; henceforth referred to as $BOADICEA_{FH}$. For controls with no available pedigree, $BOADICEA_{FH}$ was imputed based on the distribution of $BOADICEA_{FH}$ (normally distributed with mean=0 and SD=1).

Breast cancer lifetime risk

As all cases had developed BC, lifetime risks for developing a first breast tumour were calculated for all included individuals with the BOADICEA model[16], simulating an individual to be aged one year and unaffected. Initial lifetime risks ($BOADICEA_{ILR}$) were calculated based on *BRCA* status (all negative), pedigree information (for cases) as described above, and birth year. For individuals on whom information regarding PVs in the BC genes *CHEK2, PALB2*, and *ATM* was available, initial risks included the PV carrier status of these genes as well. The initial lifetime risks were compared with the lifetime risks calculated with the above information and the PRS_{313} ($BOADICEA_{PRS313}$).

Statistical analysis

The BC lifetime risks for cases and controls with ($BOADICEA_{PRS313}$) and without ($BOADICEA_{ILR}$) inclusion of the PRS_{313} were compared to define the change in risk category and thus advice for BC surveillance according to three different guidelines, NICE[24], NCCN[23] and IKNL[21].

To define how much of the variance in the PRS_{313} is explained by family history in this study, the degree of correlation between the standardised PRS_{313} and the $BOADICEA_{FH}$ for cases was determined by the Pearson correlation coefficient. This coefficient was calculated as well to estimate the linear correlation between the PRS_{313} of the proband (i.e. youngest BC diagnosis) and the PRS_{313} of other affected family members. If more than two family members were included, the average PRS_{313} of the family members was used. The association between overall BC (first breast tumour, invasive or *in* situ) and the PRS_{313} was determined with logistic regression using generalised estimating equations (GEE), adjusting for age and family history ($BOADICEA_{FH}$). Standard errors were corrected to account for relatedness of individuals using a robust estimator of the variance. To

reduce overfitting, association analyses included only cases that were not part of the development dataset for the PRS_{313} as described in Mavaddat et al.[5]

In a secondary analysis, we determined the association of the PRS_{313} with invasive and *in situ* BC risk separately. Cases that developed an invasive BC after the development of an *in situ* BC were only included in the invasive BC analysis with the age of diagnosis of the invasive breast tumour. Two of these cases were excluded because their age of diagnosis of invasive breast tumour was unknown.

In addition, the association between BC risk and the prevalence of a truncating variant in each of the 34 genes included in the BRIDGES gene panel[29] was determined with a two-sided Fisher Exact test.

Statistical significance was established at 5%. Analysis was performed using R version 4.0.3[38].

Results

The analyses included 3,918 cases from 3,492 families and 3,474 female population controls. In the association analyses, a subset of cases were included, i.e., those not included previously in the development dataset of the PRS_{313}[5]. These comprised 1,968 cases from 1,602 families (Figure S1, Table 1).

Characteristics of the included cases and controls are shown in Table 1. The mean age at last follow up for controls and age at diagnosis for cases was similar, 45 years, with an age range between 18 and 93 years. Most of the included cases had an invasive breast tumour (91%), 8% an *in situ* breast tumour and 1% a tumour of unknown invasiveness. Of all included cases, 18% developed a second breast tumour. The standardised PRS_{313} was higher for cases compared with controls with a mean of 0.71 (SD=0.96) compared with 0 for controls (SD=1.03). Distribution curves and descriptives of the standardised PRS_{313}, ER-positive PRS_{313}, and ER-negative PRS_{313} are shown in Figures S2 and S3 and Tables S2 and S3. In total, 218 (8.4%) cases and 47 (1.8%) controls were carriers of a truncating PV in either *ATM, CHEK2* or *PALB2,* excluding PVs in the last exon.

Table 1. Characteristics of participants

		Population controls	Family-based cases	Family-based cases – subset[a]
N		3,474	3,918	1,968
Families			3,492	1,602
Relatives per family included	**1**	3,474	3,099	1,263
	2	0	364	309
	3	0	25	25
	4	0	4	3
Study	**ABCS**	1,563	904	82
	HEBON	0	2,248	1,671
	ORIGO	987	0	0
	RBCS	924	766	215
Array	**GSA**		1,781	1,781
	iCOGS	2,388	1,680	163
	OncoArray	1,086	457	24
Age	**Mean**	45,6	45,1	46,8
	Range	18-93	21-91	21-91
First breast cancer	**Invasive**	NA	3,575	1,630
	In situ	NA	312	308
	Unknown	NA	31	30
ER status	**Positive**	NA	1,755	927
	Negative	NA	488	213
	Unknown	NA	1,675	828
Second breast tumour (N)		NA	719	327
Age	**Mean**	NA	52.6	52.9
	Range	NA	26-80	26-79
	Unknown	NA	130	29
Invasiveness	**Invasive**	NA	460	220
	In situ	NA	116	77
	Unknown	NA	144	30
ER status	**Positive**	NA	290	153
	Negative	NA	49	21
	Unknown	NA	380	153
Gene panel results	**All**	2,584	2,586	1,586
	No PV	2,537	2,369	1,463
	CHEK2 PV	31	167	98
	ATM PV	9	39	18
	CHEK2+ATM PV	0	2	1
	PALB2 PV	7	10	6
Standardised PRS_{313} (SD)	**Overall BC**	0 (1.03)	0.71 (0.96)	0.64 (0.88)
	ER+ BC	0 (1.03)	0.72 (0.97)	0.65 (0.88)
	ER- BC	0 (1.01)	0.45 (0.94)	0.29 (0.85)
$BOADICEA_{FH}$	**Mean (SD)**	0 (0.99)	0.55 (0.39)	0.69 (0.35)
Affected FDR	**0**	NA	1,125	
	1	NA	1,454	
	2	NA	555	
	>2	NA	176	
Affected SDR	**0**	NA	1,360	
	1	NA	1086	
	2	NA	583	
	>2	NA	281	
	Unknown	NA	615	

[a]Cases included in the association analyses which were not part of the development dataset for the PRS_{313} as described in Mavaddat et al.[5]
Abbreviations: $BOADICEA_{FH}$, Polygenic Load in calculated in the Breast and Ovarian Analysis of Disease Incidence and Carrier Estimation Algorithm; FDR, First Degree Relatives; N, Number of individuals; PRS, Polygenic Risk Score; PV, Pathogenic Variant; SD, Standard Deviation; SDR, Second Degree Relatives

Gene panel results

The BRIDGES study[29] completed sequencing for 2,037 cases with clinical data and 2,584 controls. Truncating (likely) PVs were found in 22 of 34 genes for 227 (11.1%) cases and 105 (4.1%) controls (Table S4). The majority (6.4% of the cases; 1.2% of the controls) had a truncating variant in *CHEK2,* nearly all the founder PV c.1100delC. In addition, truncating variants were relatively frequently found in *ATM, FANCM* and *PALB2* (1.8%, 0.7%, 0.6% of the cases and 0.3%, 0.6% and 0.3% of the controls respectively). The number of (pathogenic) missense variants are listed in Table S5.

PRS-based individualised risk score

Adding the PRS_{313} into the BOADICEA model ($BOADICEA_{PRS313}$) changed the absolute lifetime risk for almost all women (Figure 1), up to 34.5% for cases and up to 22.1% for controls (Figure S4, and Table S6). Clinically relevant shifts, i.e. from one to another screening category, as based on the IKNL[21], NICE[24], or NCCN[23] guidelines, were 32.4%, 36.0%, and 25.7% respectively for 1,331 cases without a gene test-result (i.e. only tested negative for a BRCA1/2 PV in diagnostic setting) (Tables 2, S7, S8). Similar results were seen for 2,369 cases that were known non-carrier of a PV in *PALB2*, *CHEK2* and *ATM*. In both groups and all age categories, a higher percentage of cases shifted to the moderate and high-risk category compared to the low-risk category (Table S9). Change towards higher risk categories was less frequent in controls than in cases (Tables S7 and S8). For cases carrying a PV in *ATM* or *CHEK2*, the proportions changing risk category were 26.3% and 17.9%, respectively, for IKNL, and 23.4% and 17.9% for NICE guidelines, but substantially lower based on the NCCN guideline (6.7% and 0.0%); this was due to the single cut-off point of 20% in the NCCN guideline. The 10 *PALB2* PV carriers in the study did not change risk category for either three guidelines.

Of the 890 controls without a gene-test result for *ATM, CHEK2*, or *PALB2* status, 4.4%, 12.0%, and 4.4% changed to another risk category based on the IKNL, NICE, and NCCN, guidelines respectively. Similar results were seen for the group where no PV was found. For *CHEK2* PV carriers, and to a lesser extent *ATM* PV carriers, these percentages were higher. Similar to cases, no change in risk category was seen for the 7 controls with a *PALB2* PV, carriers with either of three guidelines.

The distributions of the absolute lifetime risk after including the PRS_{313} for all groups ($BOADICEA_{prs313}$) are shown in Figure S5.

Table 2. Breast cancer lifetime risk category change based on the IKNL guideline

Group	BOADICEA Lifetime risk		No gene-test result		Non-PV carriers		*CHEK2* PV carriers[a]		*ATM* PV carriers[a]		*PALB2* PV carriers	
	without PRS_{313}	Including PRS_{313}	N	% change	N	% change	N	% change	N	% change	N	% change
Cases	**<20%**	**<20%**	697	30.4	1,126	30.1	3	70.0	NA		NA	
		>20%	305		486		7					
		20-30%	161	42.5	376	43.5	27	52.6	0	100.0	NA	
	20-30%	**<20%**	37		149		4		0			
		>30%	82		141		26		5			
	>30%	**>30%**	42	14.3	65	28.6	93	7.0	32	5.9	10	0.0
		<30%	7		26		7		2		0	
		Overall change		**32.4**		**33.9**		**26.3**		**17.9**		**0.0**
Controls	**<20%**	**<20%**	851	4.4	2,429	4.7	NA		NA		NA	
		>20%	39		118							
		20-30%	NA		NA		13	58.1	4	55.6	NA	
	20-30%	**<20%**					12		1			
		>30%					6		4			
	>30%	**>30%**	NA		NA		NA		NA		7	0.0
		<30%									0	
		Overall change		**4.4**		**4.7**		**58.1**		**55.6**		**0.0**

[a]Two individuals with both a pathogenic variant in *CHEK2* and *ATM* were excluded.

In total, 1,331 cases and 890 controls were included without a gene-test result; 2,369 cases and 2,537 controls in the non-PV carrier group; 167 cases and 31 controls in the CHEK2 PV carrier group; 39 cases and 9 controls in the ATM carrier group; 10 cases and 7 controls in the PALB2 PV carrier group. Abbreviations: BOADICEA, the Breast and Ovarian Analysis of Disease Incidence and Carrier Estimation Algorithm; IKNL, Netherlands Comprehensive Cancer Organisation guideline; PRS, Polygenic Risk Score; PV, Pathogenic Variant.

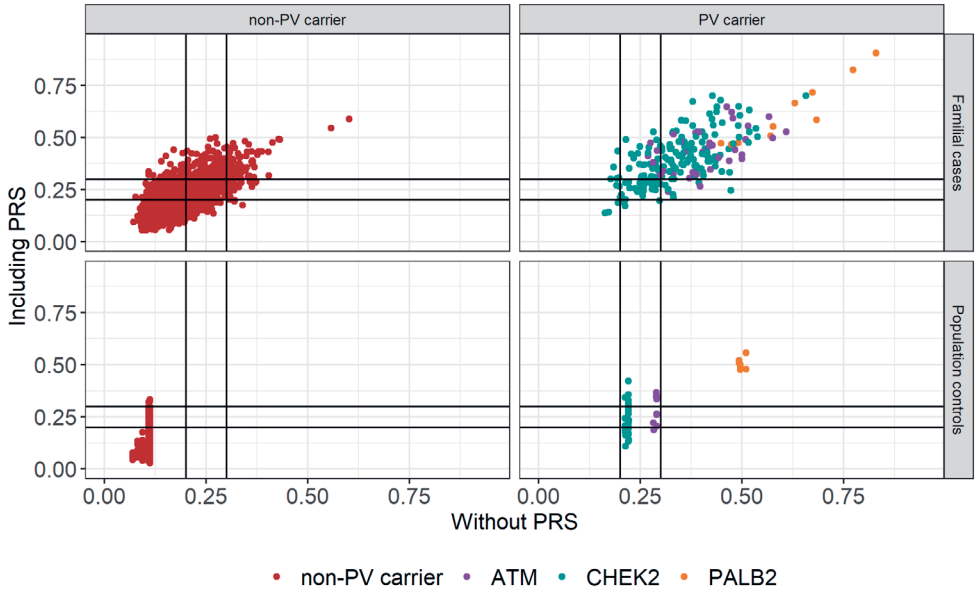

Figure 1. Change in individual breast cancer lifetime risk after including the PRS_{313}
Scatter plot of the change in breast cancer lifetime risk. For every individual, $BOADICEA_{ILR}$ was plotted against $BOADICEA_{PRS313}$. Non-carriers do not have a pathogenic variant in *ATM, CHEK2* or *PALB2* in addition to *BRCA1/2*. The solid lines represent the 20% and 30% breast cancer lifetime risk cut-off levels based on the Dutch IKNL breast cancer screening guideline[21].
Abbreviations: $BOADICEA_{ILR}$, initial breast cancer lifetime risk at age 80, based on *BRCA* status (all negative), *CHEK2*, *ATM* and *PALB2* status (if applicable), pedigree information (for cases), and birth year. $BOADICEA_{PRS313}$, breast cancer lifetime risk at age 80 including the PRS_{313} in addition to initial breast cancer lifetime risk; PRS, Polygenic Risk Score.

Table 3: Results of the association analyses between breast cancer and the PRS_{313}

		N (cases)	OR	95% CI	P-value
Main analysis	**Overall breast cancer**	1,968	1.97	1.84-2.11	$<2.00x10^{-16}$
Secondary analyses[a]	**Invasive breast cancer**	1,701	2.00	1.86-2.15	$<2.00x10^{-16}$
	***In situ* breast cancer**	262	1.69	1.50-1.89	$<2.00x10^{-16}$
Categorical PRS_{313}[b]	**0-10**	21	0.10	0.06-0.17	$<2.00x10^{-16}$
	10-20	58	0.30	0.21-0.42	$2.30x10^{-11}$
	20-40	222	0.66	0.52-0.82	$2.20x10^{-04}$
	40-60 [reference]	354	1.00	NA	NA
	60-80	491	1.37	1.13-1.66	$1.10x10^{-3}$
	80-90	396	2.27	1.84-2.79	$1.10x10^{-14}$
	90-100	426	2.29	1.86-2.83	$8.90x10^{-15}$

[a]Individuals with unknown invasiveness (N=3) and individuals with unknown age of diagnosis of the (second) invasive breast tumour (N=2) were excluded.
[b]Catagory boundaries of the PRS_{313} were -3.93; -1.27; -0.88; -0.26; 0.23; 0.84; 1.34; 3.41.
Abbreviations: CI, Confidence Interval; N, Number; OR, Odds Ratio; PRS, Polygenic Risk Score.

Correlation analysis

For cases, there was a very weak correlation between the PRS_{313} and the $BOADICEA_{FH}$ (r=0.053, p-value=8.23x10^{-4}); only 0.3% of the variance in the PRS_{313} is explained by family history. This poor correlation is visualised in Figures S6 and S7, where respectively the continuous and categorical $BOADICEA_{FH}$ are shown versus the PRS_{313}.

In contrast, there was a significant correlation between the PRS_{313} of the 393 probands and that of their affected family members (r=0.333, p-value= 1.00x10^{-11}; Figure 2)

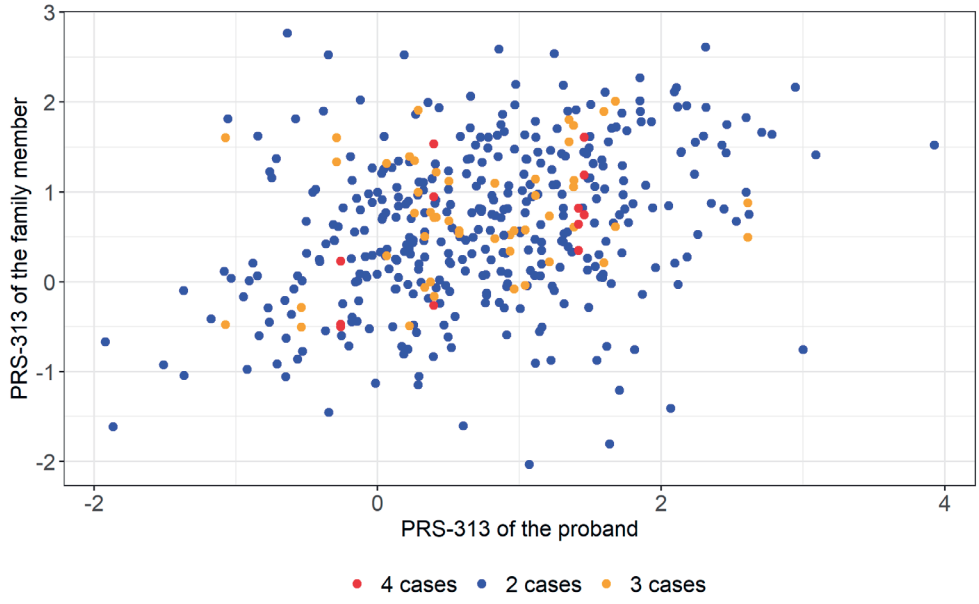

Figure 2. Correlation between the PRS_{313} of the proband and their family members
Scatter plot of the PRS_{313} of the proband (youngest breast cancer diagnosis) and their family members. Families with two individuals included are shown as blue dots, three individuals included with orange dots and four individuals included with red dots.

Abbreviations: PRS, Polygenic Risk Score.

Association analyses of PRS and breast cancer

The PRS_{313} was significantly associated with overall BC, OR per SD=1.97, 95%CI [1.84-2.11], p-value ≤2.00x10^{-16} (Table 3, Figure S8). The analyses per decile followed the trend for the continuous PRS_{313}, despite that the confidence intervals of the two lowest and the highest categories did not overlap with the continuous line (Table 3; Figure S9).

Secondary analyses for invasive BC showed similar results. *In situ* BC was also significantly associated with the PRS_{313}, OR=1.69, 95%CI [1.50-1.89], p-value ≤2.00x10^{-16} (Table 3, Figure S8).

Discussion

In this study, we have shown that the best performing PRS for BC at this moment[5] leads to substantially different patient stratification than the currently used in a familial cancer setting, which supports the implementation of the PRS_{313} in standard care for individuals from these families in clinical genetic services. Using a validated, comprehensive risk prediction model, BOADICEA[16, 39], pedigree-based family history can be easily combined with the individual PRS_{313}, as well as with gene panel results, to calculate a personal BC lifetime risk. We have shown that this procedure leads to a different risk category and corresponding clinical advice for substantial numbers of both non-carriers and carriers of a PV in a moderate BC risk gene. Furthermore, our results confirm the association between BC risk and the PRS_{313} in familial BC cases in the Dutch population[5, 40].

For *ATM* and *CHEK2* PV carriers, previous studies showed that including the PRS is of additive value for risk prediction and risk management[13, 14, 41]. A population-based study using a PRS of 105 variants[13] and a case-control study using a PRS of 86 variants[14] found similar results for *CHEK2* PV carriers and showed that there is no need for intensified breast screening for about 30% of these women. Dissimilar percentages were found for *ATM* carriers; about 50% based on the PRS-105, but a substantially lower percentage using the PRS-86 would not need intensified screening after including the PRS[13, 14]. These results were based on the NCCN guideline with a single cut-off of 20% guiding clinical management. Compared to these results and using the same guideline, we found a slightly higher percentage of *CHEK2* carriers in the unaffected population would have received different screening advice (39%), but a much lower percentage (7%) for cases with a positive family history. Although we did not see a shift in screening category for *PALB2* carriers, there was an absolute risk difference with a maximum of 9.8% for cases and 4.8% for population controls, corresponding to a lifetime risk range of 47%-91% for cases and 48%-56% for controls. A previous study found a similar effect for cases by including the PRS[42]. Such differences in risk could inform choices regarding preventive surgeries.

Our study did not have enough power to perform an association analysis between the PRS and BC for PV carriers in *PALB2, CHEK2* or *ATM*. However, previous studies showed that the per-SD effect size of a PRS with BC in PV carriers of moderate BC genes, such as *CHEK2*, is similar as in non-carriers or untested individuals[13, 43] but lower in carriers of PV in *BRCA1/2*[12]. Few studies have been performed on *ATM* or *PALB2* carriers, but a recent study showed that the effect sizes of the associations were in between those for *BRCA1/2* and CHEK2[14]. However, BOADICEA assumes that the effect of the PRS is similar for non-PV carriers and carriers of a PV in the genes *PALB2, ATM,* and *CHEK2,* i.e., pathogenic variants and the PRS contribute to risk independently. This may need some adjustment once the exact per SD effect sizes and interactions are known for these specific genes.

We found a higher effect size for the association between BC and the PRS_{313} (OR=1.97, 95%CI=1.84-2.11) than found in the population-based cohorts of BCAC (OR=1.61, 95%CI=1.57-1.65)[5] or the Dutch population (HR=1.56, 95%CI=1.40-1.73)[40]. This can possibly be explained by a higher genetic predisposition in families that visit the clinical genetic centre for counselling. Although we adjusted for family history, the weak correlation between the PRS and family history showed that adjustment for family history does not suffice to correct for the higher genetic predisposition based on the common low-risk variants. Furthermore, family history ($BOADICEA_{FH}$) for controls was imputed based on the assumption that the family history in controls was normally distributed with mean=0. This might have introduced a bias since the real family history of each control is unknown.

The virtually absent correlation between family history and the PRS_{313} was found in previous studies as well[7, 10, 18], underscoring the additive value of including the PRS in family-based risk prediction. However, to avoid double counting this requires careful joint consideration of family history and an explicelty measured PRS as provided by the BOADICEA algorithm. Altogether, the risk stratification by using the PRS in addition to family-based risk prediction in non-carriers and PV carriers highlights the need for using a comprehensive model including the PRS to calculate individual BC lifetime risks to guide screening and prevention advice. Of note, there is also no evidence that the per-SD PRS_{313} odds ratio differs across strata defined by lifestyle and hormonal risk factors[44].

Strengths of this study include the detailed family history that was available for cases. As we used only cases who visited clinical genetic centres for counselling, this cohort is a good representation of the families that are seen in a clinical genetic context. Furthermore, our results are based on a well-validated comprehensive risk prediction model, BOADICEA that has been shown to have accurate risk predictions for the general population and in familial setting[39, 40]

A limitation of this study is that we had only data for women of European ancestry, even though some studies have shown that (a subset of) the PRS_{313} is associated with BC in other ancestries as well[45, 46]. For Asian[45] and Latina[46] populations the PRS showed similar performance as in the European population, but for the African population[47] there was an attenuated effect size. Therefore, caution is needed for comprehensive risk prediction including the PRS for women of African ancestry.

In summary, including the PRS_{313} in family history-based risk prediction may change screening recommendations in up to 34% of the individuals from families with no PVs in any of the five BC genes modelled in BOADICEA. Adding the PRS_{313} also had a large impact on screening recommendations for *ATM* and *CHEK2* PV carriers. Because BOADICEA has been prospectively validated and calibrated[39, 40], clinical implementation of comprehensive

risk prediction should be considered, although this will be a logistic challenge for clinical genetic centres and would require clinical geneticists to become aware of its limitations.

Acknowledgements and funding

This work was supported by the Dutch Cancer Society (KWF), grant UL2014-7473.

We would like to thank Mary Velthuizen (University Medical Center Utrecht), J. Margriet Collée (Erasmus Medical Center Rotterdam), Wendy Prager-van den Smissen (Erasmus Medical Center Rotterdam), Daoud Ait Moha (The Netherlands Cancer Institute - Antoni van Leeuwenhoek hospital), Jan C. Oosterwijk (University Medical Center Groningen), Janet R. Vos (Radboud University Medical Center), Jeroen F.J. Laros (Leiden University Medical Center) and Adri J.M. Krol (Leiden University Medical Center) for collecting and converting the pedigree files of the included families.

Disclosure of potential conflicts of interest

AL is listed as an inventor of BOADICEA V5, which is commercialised through Cambridge Enterprise, part of Cambridge University.

References

1. Ferlay J, Colombet M, Soerjomataram I, et al. Cancer incidence and mortality patterns in Europe: Estimates for 40 countries and 25 major cancers in 2018. *European journal of cancer (Oxford, England : 1990)*. Nov 2018;103:356-387. doi:10.1016/j.ejca.2018.07.005
2. Ripping TM, Verbeek AL, Fracheboud J, de Koning HJ, van Ravesteyn NT, Broeders MJ. Overdiagnosis by mammographic screening for breast cancer studied in birth cohorts in The Netherlands. *International journal of cancer*. Aug 15 2015;137(4):921-9. doi:10.1002/ijc.29452
3. Familial breast cancer: collaborative reanalysis of individual data from 52 epidemiological studies including 58,209 women with breast cancer and 101,986 women without the disease. *Lancet (London, England)*. Oct 27 2001;358(9291):1389-99. doi:10.1016/s0140-6736(01)06524-2
4. Michailidou K, Lindstrom S, Dennis J, et al. Association analysis identifies 65 new breast cancer risk loci. *Nature*. Oct 23 2017;doi:10.1038/nature24284
5. Mavaddat N, Michailidou K, Dennis J, et al. Polygenic Risk Scores for Prediction of Breast Cancer and Breast Cancer Subtypes. *American journal of human genetics*. Jan 3 2019;104(1):21-34. doi:10.1016/j.ajhg.2018.11.002
6. Mavaddat N, Pharoah PD, Michailidou K, et al. Prediction of breast cancer risk based on profiling with common genetic variants. *JNatlCancer Inst*. 5/2015 2015;107(5)Not in File. doi:djv036 [pii];10.1093/jnci/djv036 [doi]
7. Lakeman IMM, Hilbers FS, Rodriguez-Girondo M, et al. Addition of a 161-SNP polygenic risk score to family history-based risk prediction: impact on clinical management in non-BRCA1/2 breast cancer families. *Journal of medical genetics*. Sep 2019;56(9):581-589. doi:10.1136/jmedgenet-2019-106072
8. Sawyer S, Mitchell G, McKinley J, et al. A role for common genomic variants in the assessment of familial breast cancer. *JClinOncol*. 12/10/2012 2012;30(35):4330-4336. Not in File. doi:JCO.2012.41.7469 [pii];10.1200/JCO.2012.41.7469 [doi]
9. Li H, Feng B, Miron A, et al. Breast cancer risk prediction using a polygenic risk score in the familial setting: a prospective study from the Breast Cancer Family Registry and kConFab. *Genetics in medicine : official journal of the American College of Medical Genetics*. May 12 2016;doi:10.1038/gim.2016.43
10. Muranen TA, Mavaddat N, Khan S, et al. Polygenic risk score is associated with increased disease risk in 52 Finnish breast cancer families. *Breast cancer research and treatment*. Aug 2016;158(3):463-9. doi:10.1007/s10549-016-3897-6
11. Kuchenbaecker KB, McGuffog L, Barrowdale D, et al. Evaluation of Polygenic Risk Scores for Breast and Ovarian Cancer Risk Prediction in BRCA1 and BRCA2 Mutation Carriers. *Journal of the National Cancer Institute*. Jul 01 2017;109(7)doi:10.1093/jnci/djw302
12. Barnes D, Rookus MA, McGuffog L, et al. Polygenic risk scores and breast and epithelial ovarian cancer risks for carriers of BRCA1 and BRCA2 pathogenic variants. *Genetics in Medicine*. 2020;accepted for publication

13. Gao C, Polley EC, Hart SN, et al. Risk of Breast Cancer Among Carriers of Pathogenic Variants in Breast Cancer Predisposition Genes Varies by Polygenic Risk Score. *Journal of Clinical Oncology*. 0(0):JCO.20.01992. doi:10.1200/jco.20.01992
14. Gallagher S, Hughes E, Wagner S, et al. Association of a Polygenic Risk Score With Breast Cancer Among Women Carriers of High- and Moderate-Risk Breast Cancer Genes. *JAMA Network Open*. 2020;3(7):e208501-e208501. doi:10.1001/jamanetworkopen.2020.8501
15. Cintolo-Gonzalez JA, Braun D, Blackford AL, et al. Breast cancer risk models: a comprehensive overview of existing models, validation, and clinical applications. *Breast cancer research and treatment*. Jul 2017;164(2):263-284. doi:10.1007/s10549-017-4247-z
16. Lee A, Mavaddat N, Wilcox AN, et al. BOADICEA: a comprehensive breast cancer risk prediction model incorporating genetic and nongenetic risk factors. *Genetics in medicine : official journal of the American College of Medical Genetics*. Jan 15 2019;21(8):1708-1718. doi:10.1038/s41436-018-0406-9
17. Cuzick J, Brentnall AR, Segal C, et al. Impact of a Panel of 88 Single Nucleotide Polymorphisms on the Risk of Breast Cancer in High-Risk Women: Results From Two Randomized Tamoxifen Prevention Trials. *Journal of clinical oncology : official journal of the American Society of Clinical Oncology*. Mar 2017;35(7):743-750. doi:10.1200/jco.2016.69.8944
18. Dite GS, MacInnis RJ, Bickerstaffe A, et al. Breast Cancer Risk Prediction Using Clinical Models and 77 Independent Risk-Associated SNPs for Women Aged Under 50 Years: Australian Breast Cancer Family Registry. *Cancer EpidemiolBiomarkers Prev*. 12/16/2015 2015;Not in File. doi:1055-9965.EPI-15-0838 [pii];10.1158/1055-9965.EPI-15-0838 [doi]
19. Shieh Y, Hu D, Ma L, et al. Breast cancer risk prediction using a clinical risk model and polygenic risk score. *Breast cancer research and treatment*. Oct 2016;159(3):513-25. doi:10.1007/s10549-016-3953-2
20. van Veen EM, Brentnall AR, Byers H, et al. Use of Single-Nucleotide Polymorphisms and Mammographic Density Plus Classic Risk Factors for Breast Cancer Risk Prediction. *JAMA oncology*. Apr 1 2018;4(4):476-482. doi:10.1001/jamaoncol.2017.4881
21. IKNL. Richtlijn Borstkanker - Screening buiten het bevolkingsonderzoek. Accessed 03-12-2021, https://richtlijnendatabase.nl/richtlijn/borstkanker/screening/screening_buiten_het_bob/screening_buiten_het_bevolkingsonderzoek.html
22. Carver T, Hartley S, Lee A, et al. CanRisk Tool—A Web Interface for the Prediction of Breast and Ovarian Cancer Risk and the Likelihood of Carrying Genetic Pathogenic Variants. *Cancer Epidemiology Biomarkers & Prevention*. 2021;30(3):469-473. doi:10.1158/1055-9965.Epi-20-1319
23. NCCN. Clinical Practice Guidelines in Oncology; Breast Cancer Screening and Diagnosis. 2017. Available from: https://www.nccn.org/professionals/physician_gls/pdf/breast-screening.pdf. Accessed April, 2018
24. NICE. National Institute for Health and Care Excellence: Familial breast cancer: classification, care and managing breast cancer and related risks in people with a family history of breast cancer. 2013. Available from: www.nice.org.uk/guidance/cg164. Accessed April, 2018;

25. von Elm E, Altman DG, Egger M, Pocock SJ, Gøtzsche PC, Vandenbroucke JP. Strengthening the Reporting of Observational Studies in Epidemiology (STROBE) statement: guidelines for reporting observational studies. *BMJ (Clinical research ed)*. Oct 20 2007;335(7624):806-8. doi:10.1136/bmj.39335.541782.AD
26. van Asperen CJ, Brohet RM, Meijers-Heijboer EJ, et al. Cancer risks in BRCA2 families: estimates for sites other than breast and ovary. *Journal of medical genetics*. Sep 2005;42(9):711-9. doi:10.1136/jmg.2004.028829
27. Schmidt MK, Hogervorst F, van Hien R, et al. Age- and Tumor Subtype-Specific Breast Cancer Risk Estimates for CHEK2*1100delC Carriers. *Journal of clinical oncology : official journal of the American Society of Clinical Oncology*. Aug 10 2016;34(23):2750-60. doi:10.1200/jco.2016.66.5844
28. Liu J, Prager-van der Smissen WJ, Schmidt MK, et al. Recurrent HOXB13 mutations in the Dutch population do not associate with increased breast cancer risk. *Sci Rep*. Jul 18 2016;6:30026. doi:10.1038/srep30026
29. Dorling L, Carvalho S, Allen J, et al. Breast Cancer Risk Genes - Association Analysis in More than 113,000 Women. *The New England journal of medicine*. Jan 20 2021;doi:10.1056/NEJMoa1913948
30. McLaren W, Gil L, Hunt SE, et al. The Ensembl Variant Effect Predictor. *Genome Biology*. 2016/06/06 2016;17(1):122. doi:10.1186/s13059-016-0974-4
31. Boonen R, Rodrigue A, Stoepker C, et al. Functional analysis of genetic variants in the high-risk breast cancer susceptibility gene PALB2. *Nat Commun*. Nov 22 2019;10(1):5296. doi:10.1038/s41467-019-13194-2
32. Landrum MJ, Lee JM, Benson M, et al. ClinVar: improving access to variant interpretations and supporting evidence. *Nucleic Acids Res*. Jan 4 2018;46(D1):D1062-d1067. doi:10.1093/nar/gkx1153
33. Michailidou K, Hall P, Gonzalez-Neira A, et al. Large-scale genotyping identifies 41 new loci associated with breast cancer risk. *NatGenet*. 4/2013 2013;45(4):353-2. Not in File. doi:ng.2563 [pii];10.1038/ng.2563 [doi]
34. Das S, Forer L, Schonherr S, et al. Next-generation genotype imputation service and methods. *Nature genetics*. Oct 2016;48(10):1284-1287. doi:10.1038/ng.3656
35. McCarthy S, Das S, Kretzschmar W, et al. A reference panel of 64,976 haplotypes for genotype imputation. *Nature genetics*. Oct 2016;48(10):1279-83. doi:10.1038/ng.3643
36. Deelen P, Menelaou A, van Leeuwen EM, et al. Improved imputation quality of low-frequency and rare variants in European samples using the 'Genome of The Netherlands'. *European journal of human genetics : EJHG*. Nov 2014;22(11):1321-6. doi:10.1038/ejhg.2014.19
37. Abecasis GR, Altshuler D, Auton A, et al. A map of human genome variation from population-scale sequencing. *Nature*. Oct 28 2010;467(7319):1061-73. doi:10.1038/nature09534
38. R_Core_Team_(2017). R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. https://www.R-project.org/
39. Pal Choudhury P, Brook MN, Hurson AN, et al. Comparative validation of the BOADICEA and Tyrer-Cuzick breast cancer risk models incorporating classical risk factors and polygenic risk in a population-based prospective cohort of women of European ancestry. *Breast Cancer Research*. 2021/02/15 2021;23(1):22. doi:10.1186/s13058-021-01399-7

40. Lakeman IMM, Rodríguez-Girondo M, Lee A, et al. Validation of the BOADICEA model and a 313-variant polygenic risk score for breast cancer risk prediction in a Dutch prospective cohort. *Genetics in medicine : official journal of the American College of Medical Genetics*. Nov 2020;22(11):1803-1811. doi:10.1038/s41436-020-0884-4
41. Borde J, Ernst C, Wappenschmidt B, et al. Performance of Breast Cancer Polygenic Risk Scores in 760 Female CHEK2 Germline Mutation Carriers. *Journal of the National Cancer Institute*. Jul 1 2021;113(7):893-899. doi:10.1093/jnci/djaa203
42. Mars N, Widén E, Kerminen S, et al. The role of polygenic risk and susceptibility genes in breast cancer over the course of life. *Nat Commun*. Dec 14 2020;11(1):6383. doi:10.1038/s41467-020-19966-5
43. Muranen TA, Greco D, Blomqvist C, et al. Genetic modifiers of CHEK2*1100delC-associated breast cancer risk. *Genetics in medicine : official journal of the American College of Medical Genetics*. Oct 06 2016;doi:10.1038/gim.2016.147
44. Kapoor PM, Mavaddat N, Choudhury PP, et al. Combined Associations of a Polygenic Risk Score and Classical Risk Factors With Breast Cancer Risk. *Journal of the National Cancer Institute*. Mar 1 2021;113(3):329-337. doi:10.1093/jnci/djaa056
45. Ho WK, Tan MM, Mavaddat N, et al. European polygenic risk score for prediction of breast cancer shows similar performance in Asian women. *Nat Commun*. Jul 31 2020;11(1):3833. doi:10.1038/s41467-020-17680-w
46. Shieh Y, Fejerman L, Lott PC, et al. A polygenic risk score for breast cancer in U.S. Latinas and Latin-American women. *Journal of the National Cancer Institute*. Sep 25 2019;doi:10.1093/jnci/djz174
47. Du Z, Gao G, Adedokun B, et al. Evaluating Polygenic Risk Scores for Breast Cancer in Women of African Ancestry. *Journal of the National Cancer Institute*. Mar 26 2021;doi:10.1093/jnci/djab050

Supplementary methods

Study cohorts

HEBON

The HEBON study[1] (initiated in 1999) is an ongoing nationwide retrospective cohort study among breast cancer families with prospective follow up. Participants were invited after visiting one of the Clinical Genetic Centers in the Netherlands for breast and/or ovarian cancer counselling. Participants were asked to fill in a questionnaire about lifestyle, family history and risk factors for breast cancer. Linkage with the nationwide cancer and pathology registries is possible for follow up.

Additional selection criteria for HEBON participants included:

- At least two breast cancer cases in a family with available DNA samples
- Breast cancer diagnosis below the age of 60 years and a positive family history:
    - One first degree family member with breast cancer diagnosis below the age of 50 OR
    - Two first or second-degree family members with breast cancer diagnosis below the age of 60

ABCS-F and RBCS

The ABCS-F[2] and RBCS[3] case-cohorts included also breast cancer cases who visited the Clinical Genetic Centres of the Netherlands Cancer Institute in Amsterdam or the Erasmus Medical Center in Rotterdam, respectively. No additional selection criteria were used for ABCS-F and RBCS cases. 151 individuals from the ABCS-F study and 469 individuals from the RBCS study are included in the HEBON study as well and shown as HEBON cases in Table 1.

Quality control procedure

For the 2,179 breast cancer cases without a *BRCA1/2* pathogenic variant that were genotyped with the GSA array, quality control was performed with Plink version 1.9, which excluded 8,408 SNPs with a call rate below 95%. Another 712 SNPs were removed because of a deviation from Hardy-Weinberg equilibrium in controls at $P<1x10^{-12}$. In total, 124 individuals were excluded of which 62 individuals with a call rate below 95%, 7 individuals because they were genotypically not female or the gender was uncertain, and 17 individuals because of a sample swab. After population stratification analysis, 28 individuals were excluded because of non-European genotype (>3 SD).

Imputation pedigrees

In total, 3,492 pedigrees were collected for this study. These pedigrees consisted of 202,680 individuals (49% female) of which 12,785 individuals were affected with breast cancer.

If the age of breast cancer diagnosis for a family member was not known (n=1,272), a conditional average age was estimated given the age at last follow up of the individual and the breast cancer incidence in the Netherlands. Furthermore, for all affected individuals with breast cancer, ovarian cancer, prostate cancer or pancreatic cancer the year of birth was imputed, if this was not yet available, based on the year of birth of the closest relative (25 year difference for parents and children, average for siblings). If the age of last follow up was not known, this age was calculated based on the date of the last update of a pedigree and the year of birth.

Supplementary figures and tables

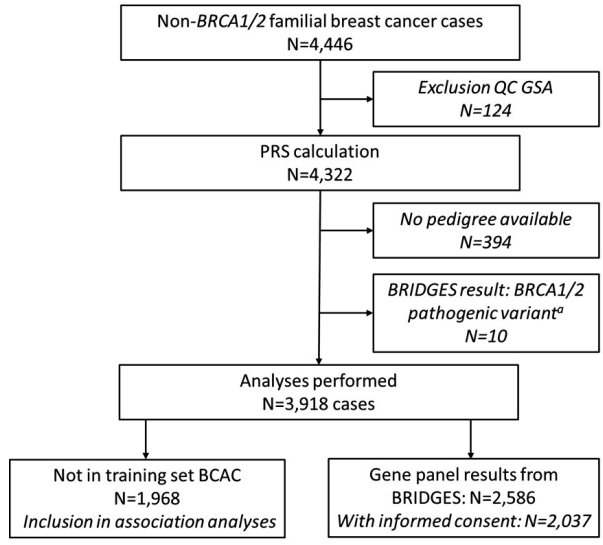

Figure S1: Flow scheme of the selection procedure
Breast cancer cases were selected from the ABCS, HEBON and RBCS studies. Details of the quality control procedure are described above. Absolute lifetime risks were calculated for all included cases (N=3,918). To exclude overlap of cases with the development dataset for the PRS_{313}[4], only 1,968 cases were included in the association analyses. For the majority of cases gene panel information was available. For cases of whom we did not have informed consent to report the clinical relevant results, only pseudoanonymized information about pathogenic variants in *ATM, CHEK2*, and *PALB2* was available (N=549). For the cases with informed consent, the number of pathogenic variants and missense variants are shown in Table S3.
[a]carriers of a pathogenic variant or family member of a carrier of a pathogenic variant in *BRCA1* or *BRCA2*.
Abbreviations: BCAC, Breast Cancer Association Consortium; BRIDGES, Breast cancer Risk after Diagnostic GEne Sequencing; PRS, Polygenic Risk Score.

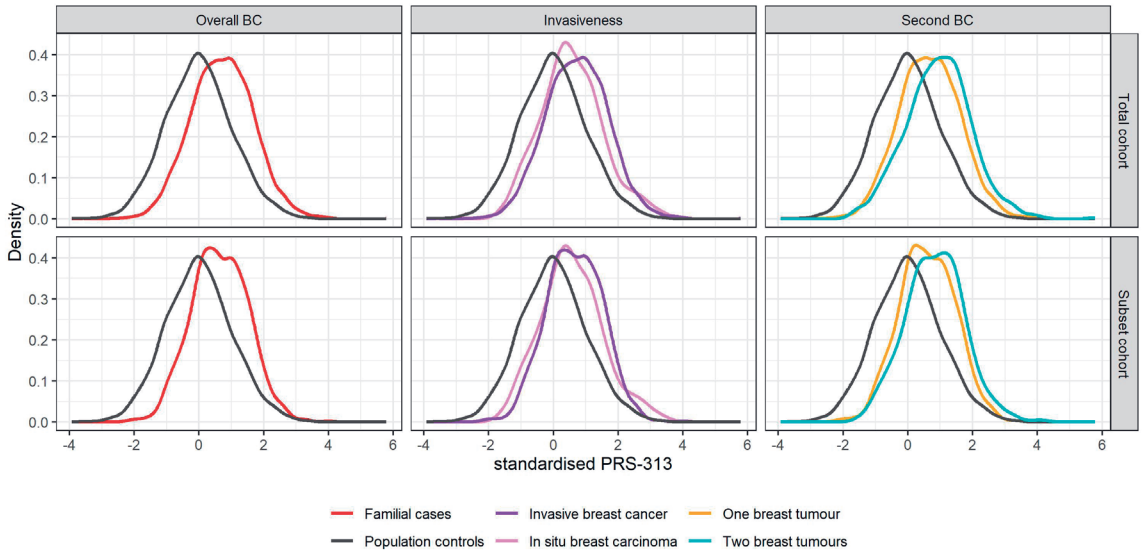

Figure S2: Density curves of the PRS_{313}
Distribution of the PRS_{313} in the included 3,474 population controls (grey line) and 3,918 and 1,968 breast cancer cases (red line) in the total and subset cohort respectively. For the invasiveness figure, 3 cases were excluded for which invasiveness for the first and/or second breast tumour was unknown. In the total cohort 3,653 and 262 cases were included with invasive (purple line) and in situ (pink line) breast cancer respectively. For the subset cohort this was 1,703 and 262. In the right figure, 719 and 327 breast cancer cases with a second breast tumour (blue line) were included in the total and subset cohort respectively.
Abbreviations: BC, Breast Cancer; PRS, Polygenic Risk Score.

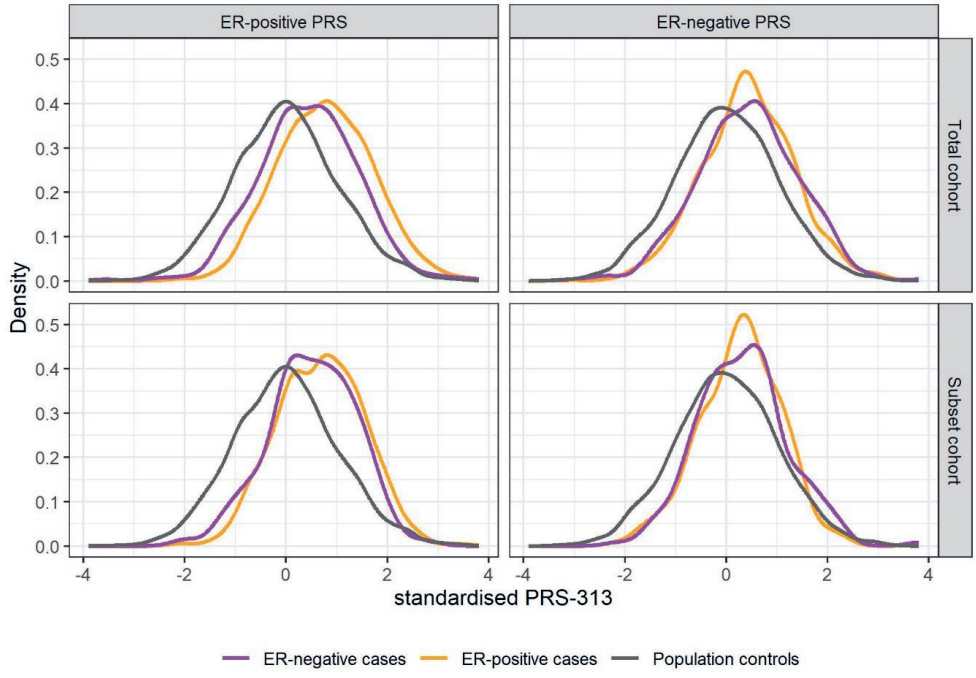

Figure S3: Density curves of the ER-positive and ER-negative PRS_{313}
Distribution of the ER-negative (left figures) and ER-positive (right figures) PRS_{313} for cases with an ER-negative (purple line) and ER-positive (orange line) first breast tumour. As a reference, the distribution of these PRS in population controls are shown as well (grey line). In the total cohort, 1,755 and 488 breast cancer cases are included with a first ER-positve and ER-negative breast tumour respectively. For the subset cohort this was 927 and 213 respectively.
Abbreviations: ER, Estrogen Receptor; PRS, Polygenic Risk Score

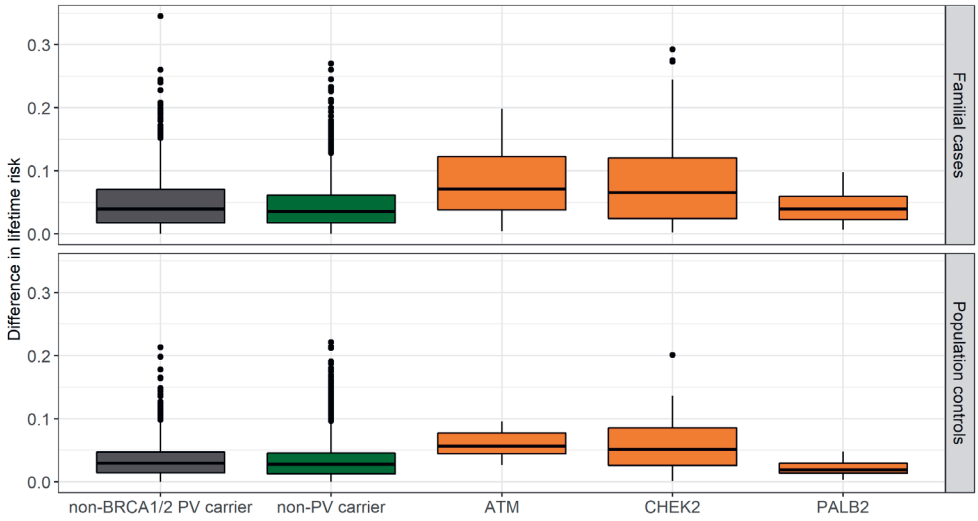

Figure S4: Difference in breast cancer lifetime risk score calculated by BOADICEA
Boxplot of the difference in breast cancer lifetime risk between the basic calculation in BOADICEA and after including the PRS_{313}. The basic calculation included birth year, gene panel results and for cases a pedigree of their family in addition. Non-carriers are the group of which we know that they do not have a pathogenic variant in *ATM, CHEK2* and *PALB2* in addition to *BRCA1/2.*
Abbreviations: BOADICEA, Breast and Ovarian Analysis of Disease Incidence and Carrier Estimation Algorithm; PV, Pathogenic Variant.

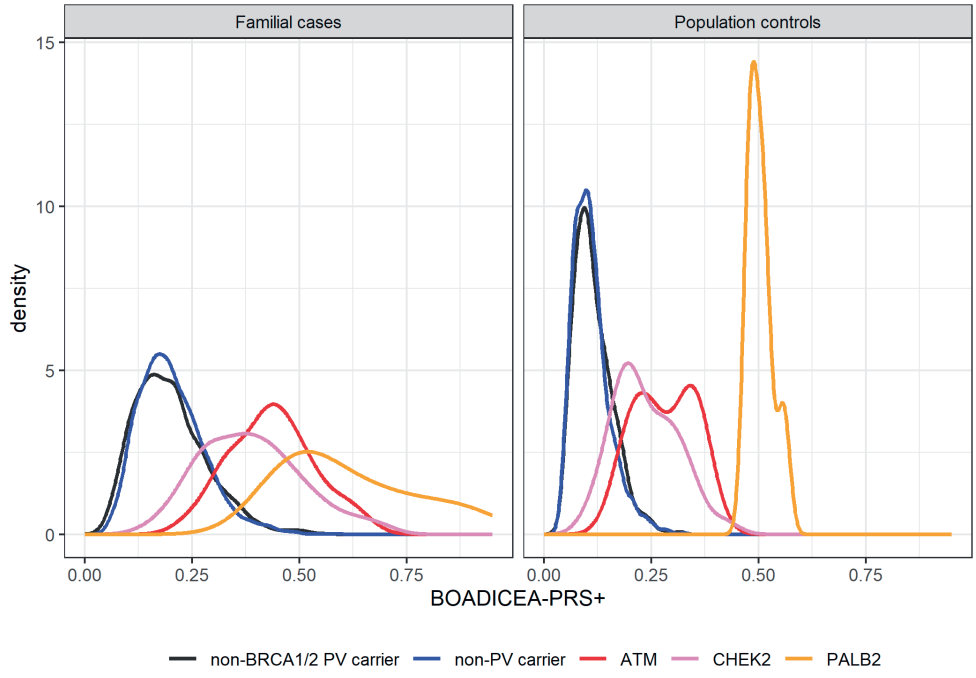

Figure S5. Distribution of breast cancer lifetime risk after including the PRS_{313}
Density plots of the distribution in breast cancer lifetime risk calculated with BOADICEA including birth cohort, gene panel results, pedigree-based family history for cases and the PRS_{313}.
Abbreviations: BOADICEA, Breast and Ovarian Analysis of Disease Incidence and Carrier Estimation Algorithm; PV, Pathogenic Variant; PRS, Polygenic Risk Score

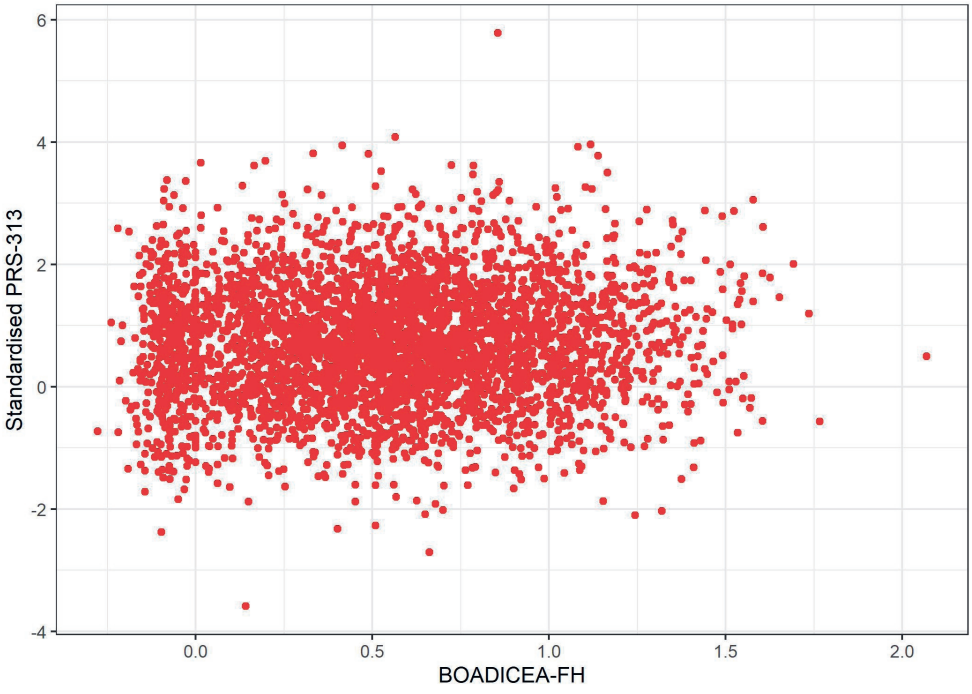

Figure S6. Correlation plot between de $BOADICEA_{FH}$ and the PRS_{313}
For all included breast cancer cases (N=3,918), the individual $BOADICEA_{FH}$ (polygenic load) is plotted against the PRS_{313}. $BOADICEA_{FH}$ was calculated with BOADICEA based on the pedigree without inclusion of the PRS_{313}.
Abbreviations: BOADICEA, Breast and Ovarian Analysis of Disease Incidence and Carrier Estimation Algorithm; FH, Family History; PRS, Polygenic Risk Score.

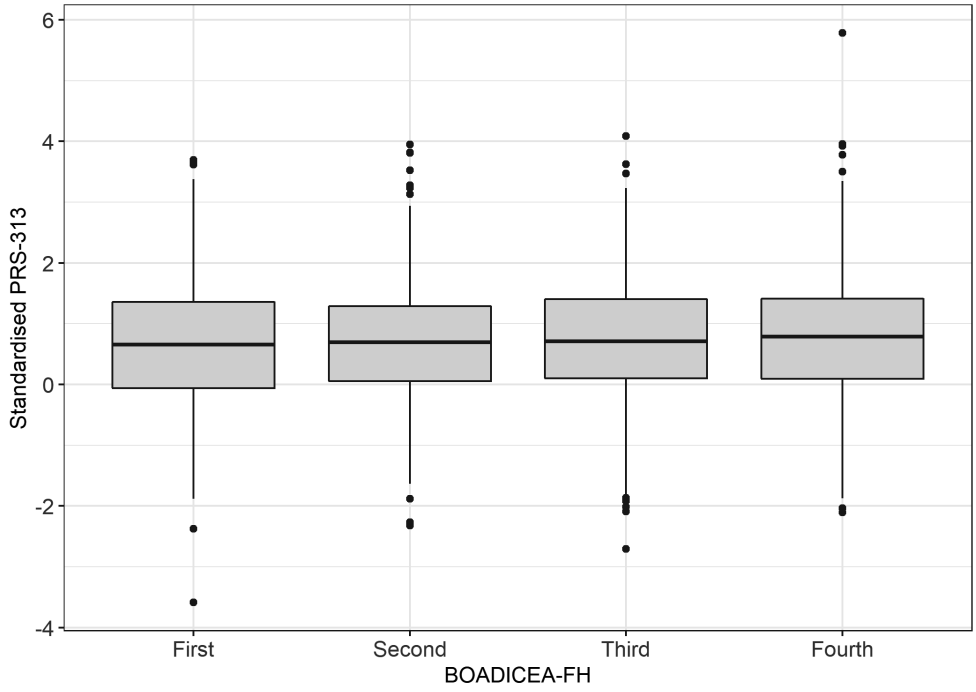

Figure S7: PRS_{313} distribution by quartiles of $BOADICEA_{FH}$
The PRS_{313} distribution for all included cases (N=3,918) separated by quartiles of the individual $BOADICEA_{FH}$ (polygenic load). $BOADICEA_{FH}$ was calculated with BOADICEA based on the pedigree without inclusion of the PRS_{313}.
Abbreviations: BOADICEA, Breast and Ovarian Analysis of Disease Incidence and Carrier Estimation Algorithm; FH, Family History; PRS, Polygenic Risk Score.

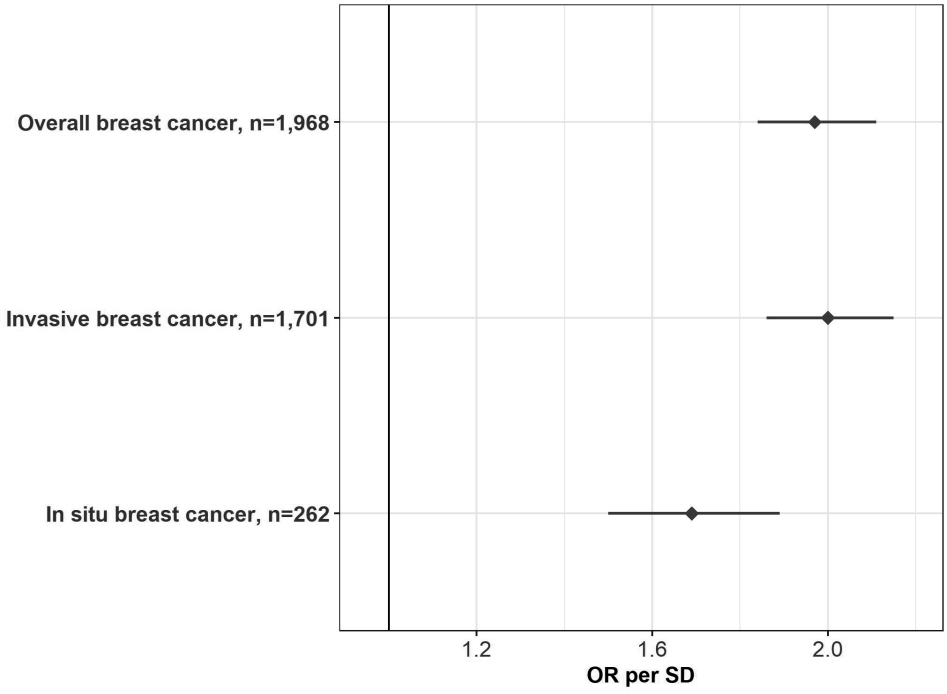

Figure S8: Association between the PRS_{313} and breast cancer
Visualisation of the effect sizes and 95% confidence intervals of the association between the PRS_{313} and breast cancer. The corresponding OR and included breast cancer cases are shown in Table 3. Abbreviations: BC, Breast Cancer; OR, Odds Ratio; PRS, Polygenic Risk Score

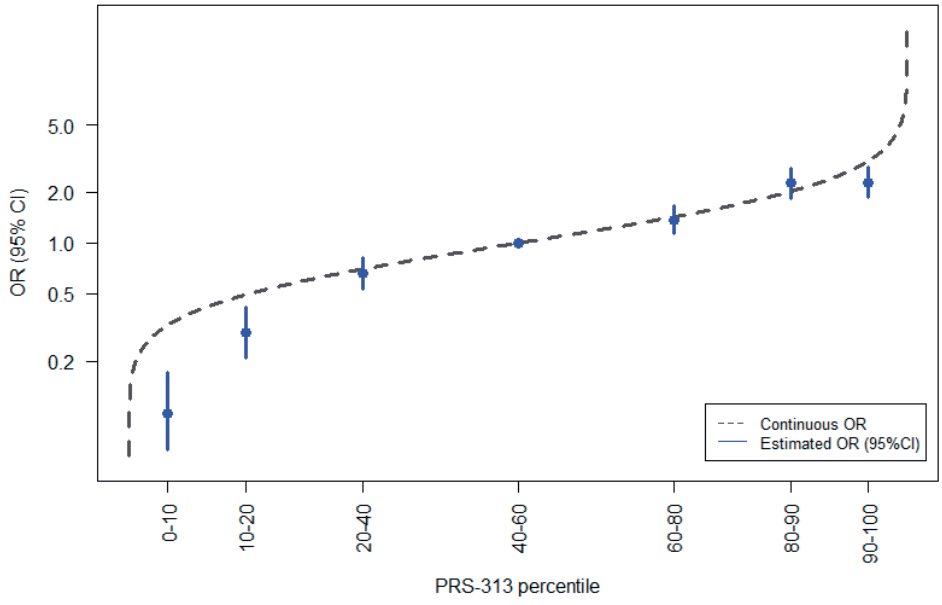

Figure S9: Association between the PRS and breast cancer by percentiles of the PRS_{313}
Plot of the effect size of the association between the continuous PRS_{313} (grey line) and breast cancer and the categorical PRS_{313} (blue dots) and breast cancer. Corresponding OR and 95% confidence intervals are shown in Table 3.
Abbreviations: CI, Confidence Interval; OR, Odds Ratio; PRS, Polygenic Risk Score.

Table S1: common low risk variants included in the PRS_{313} (large Excel file)
Available upon request / see online material. This table is partly published before by Mavaddat et al.[4] We added the imputation quality in this study.

Table S2: Descriptives of the standardised PRS_{313}

	Total cohort			Family-based cases – subset[c]		
	N	Mean PRS_{313}	SD PRS_{313}	N	Mean PRS_{313}	SD PRS_{313}
All cases	3,918	0.71	0.96	1,968	0.64	0.88
Invasive cases[a]	3,653	0.73	0.96	1,703	0.65	0.86
***In situ* only cases[b]**	262	0.56	0.96	262	0.56	0.96
1 breast tumour	3,199	0.66	0.95	1,641	0.60	0.87
2 breast tumours	719	0.95	1.01	327	0.83	0.90
Population controls	3,474	0	1.03	NA	NA	NA

[a]Invasive first or second tumour
[b]no invasive first or second tumour
[c]Cases included in the association analyses which were not part of the development dataset for the PRS_{313} as described in Mavaddat et al.[4]
Abbreviations: N, Number; NA, Not Applicable; PRS, Polygenic Risk Score

Table S3: Descriptives of the standardised ER-positive and ER-negative PRS_{313}

Group	PRS	Total cohort			Family-based cases – subset[c]		
		N	Mean PRS	SD PRS	N	Mean PRS	SD PRS
ER-positive BC	ER-positive PRS	1,755	0.78	0.92	927	0.68	0.86
ER-negative BC	ER-positive PRS	488	0.43	0.98	213	0.51	0.85
ER-positive BC	ER-negative PRS	1,755	0.76	0.93	927	0.66	0.85
ER-negative BC	ER-negative PRS	488	0.46	0.97	213	0.52	0.85

[a]Invasive first or second tumour
[b]no invasive first or second tumour
[c]Cases included in the association analyses which were not part of the development dataset for the PRS_{313} as described in Mavaddat et al.[4]
Abbreviations: N, Number; NA, Not Applicable; PRS, Polygenic Risk Score

Table S4: Truncating variants in BRIDGES gene panel

Gene	Cases, N=2,037[a]		Controls, N=2,584[a]		OR	95% CI	P-value
	N	%	N	%			
ABRAXAS1	1	0.0	0	0.0	NA	NA	NA
AKT1	0	0.0	0	0.0	NA	NA	NA
ATM	**36**	**1.8**	**9**	0.3	**5.15**	**2.42-12.18**	**1.00x10^{-06}**
BARD1	1	0.0	1	0.0	1.27	0.02-99.55	1.00
BRCA1	NA	NA	NA	NA	NA	NA	NA
BRCA2	NA	NA	NA	NA	NA	NA	NA
BRE	0	0.0	0	0.0	NA	NA	NA
BRIP1	4	0.2	5	0.2	1.01	0.20-4.72	1.00
CDH1	0	0.0	0	0.0	NA	NA	NA
CHEK2	**131**	**6.4**	**31**	**1.2**	**5.66**	**3.78-8.70**	**<2.00x10^{-16}**
c.1100delC[b]	130		30				
Other	1						
EPCAM	0	0.0	2	0.1	NA	NA	NA
FANCC	5	0.2	8	0.3	0.79	0.20-2.75	0.80
FANCM	14	0.7	16	0.6	1.11	0.50-2.44	0.90
GEN1	0	0.0	0	0.0	NA	NA	NA
MEN1	0	0.0	0	0.0	NA	NA	NA
MLH1	0	0.0	0	0.0	NA	NA	NA
MRE11A	1	0.0	3	0.1	0.42	0.01-5.27	0.60
MSH2	0	0.0	2	0.1	NA	NA	NA
MSH6	1	0.0	0	0.0	NA	NA	NA
MUTYH	3	0.1	2	0.1	1.9	0.22-22.81	0.70
NBN	2	0.1	3	0.1	0.85	0.07-7.39	1,00
NF1	2	0.1	0	0.0	NA	NA	NA
PALB2	12[c]	0.6	7	0.3	2.18	0.79-6.55	0.10
PIK3CA	0	0.0	0	0.0	NA	NA	NA
PMS2	1	0.0	2	0.1	0.63	0.01-12.19	1.00
PTEN	1	0.0	1	0.0	1.27	0.02-99.55	1.00
RAD50	4	0.2	7	0.3	0.72	0.16-2.85	0.80
RAD51C	1	0.0	0	0.0	NA	NA	NA
RAD51D	5	0.2	0	0.0	NA	NA	NA
RECQL	2	0.1	3	0.1	0.85	0.07-7.39	1.00
RINT1	0	0.0	2	0.1	NA	NA	NA
STK11	0	0.0	0	0.0	NA	NA	NA
TP53	0	0.0	0	0.0	NA	NA	NA
XRCC2	0	0.0	1	0.0	NA	NA	NA
Total	227	11.1	105	4.1	-	-	-

[a]Cases and controls were included in the analyses described by Dorling et al.[5]
[b]of which 6 homozygous in cases and 1 homozygous in controls
[c]In addition to inclusion criteria for truncating variants in BRIDGES, 4 *PALB2* truncating variants in the last exon were added.
Abbreviations: CI, Confidence Interval; N, Number; NA, Not Applicable; OR, Odds Ratio.

Table S5: Missense variants in BRIDGES gene panel

Gene	Cases; N=2,038[a]		Controls, N=2,584[a]	
	Total[b]	P/LP[c]	Total[b]	P/LP[c]
ABRAXAS1	3	NA	5	NA
AKT1	2	NA	6	NA
ATM	121	5	113	4
BARD1	25	0	26	0
BRCA1	42	NA	49	NA
BRCA2	109	NA	127	NA
BRE	0	NA	0	NA
BRIP1	34	NA	41	NA
CDH1	26	NA	28	NA
CHEK2	64	8	34	2
EPCAM	9	NA	18	NA
FANCC	28	NA	23	NA
FANCM	64	NA	62	NA
GEN1	38	NA	32	NA
MEN1	4	NA	2	NA
MLH1	19	NA	21	NA
MRE11A	16	NA	19	NA
MSH2	42	NA	56	NA
MSH6	51	NA	52	NA
MUTYH	28	NA	33	NA
NBN	35	NA	23	NA
NF1	30	NA	34	NA
PALB2	23	0	23	0
PIK3CA	6	NA	10	NA
PMS2	37	NA	28	NA
PTEN	3	NA	7	NA
RAD50	50	NA	46	NA
RAD51C	9	1	9	0
RAD51D	6	0	10	0
RECQL	16	NA	20	NA
RINT1	39	NA	47	NA
STK11	0	NA	1	NA
TP53	14	4	10	0
XRCC2	6	NA	13	NA
Total	999	18	1,028	6

[a]Cases and controls were included in the analyses described by Dorling et al.[5]
[b]Total number of missense variants detected, not corrected for individuals who carry more than one missense variant in a single gene.
[c]For genes in which pathogenic variants are associated with breast cancer[5], missense variant interpretation was performed by using the ClinVar database[6].
Abbreviations: N, Number; NA, Not Applicable; P, Pathogenic; LP, Likely Pathogenic.

Table S6: Absolute change in breast cancer lifetime risk after including the PRS_{313}

	Cases			Controls		
	Min	**Mean**	**Max**	**Min**	**Mean**	**Max**
Non-*BRCA1/2* PV carriers	0	5.0	34.5	0	3.5	21.3
Non-carriers	0	4.5	27.0	0	3.3	22.1
***ATM* PV carriers**[a]	0.4	8.0	19.8	2.6	5.9	9.6
***CHEK2* PV carriers**[a]	0.3	8.1	29.3	0.1	5.9	20.1
***PALB2* PV carriers**	0.7	4.4	9.8	0.3	2.2	4.8

[a]Two cases with both a pathogenic variant in *CHEK2* and *ATM* were excluded.
In total, 1,331 cases and 890 controls were included without a gene-test result; 2,369 cases and 2,537 controls in the non-PV carrier group; 167 cases and 31 controls in the CHEK2 PV carrier group; 39 cases and 9 controls in the ATM carrier group; 10 cases and 7 controls in the PALB2 PV carrier group.
Abbreviations: Min, Minimum; Max, Maximum; PRS, Polygenic Risk Score; PV, Pathogenic Variant.

Table S7: Breast cancer lifetime risk category change based on the NCCN guideline

Group	BOADICEA Lifetime risk		No gene-test result		Non-PV carriers		*CHEK2* PV carriers[a]		*ATM* PV carriers[a]		*PALB2* PV carriers	
	Without PRS_{313}	Including PRS_{313}	N	% change	N	% change	N	% change	N	% change	N	% change
Cases	**<20%**	**<20%**	697	30.4	1,126	30.1	3	70.0	0	0.0	0	0.0
		>20%	305		486		7		0		0	
	>20%	**>20%**	292	11.2	605	20.1	153	2.5	39	0.0	10	0.0
		<20%	37		152		4		0		0	
		Overall change	**25.7**		**26.9**		**6.6**		**0.0**		**0.0**	
Controls	**<20%**	**<20%**	851	4.4	2,419	4.7	NA		NA		NA	
		>20%	39		118							
	>20%	**>20%**	NA		NA		19	38.7	8	11.1	7	0.0
		<20%					12		1		0	
		Overall change	**4.4**		**4.7**		**38.7**		**11.1**		**0.0**	

[a]Two cases with both a pathogenic variant in *CHEK2* and *ATM* were excluded.

In total, 1,331 cases and 890 controls were included without a gene-test result (no BRCA1/2 PV); 2,369 cases and 2,537 controls in the non-PV carrier group; 167 cases and 31 controls in the CHEK2 PV carrier group; 39 cases and 9 controls in the ATM carrier group; 10 cases and 7 controls in the PALB2 PV carrier group.

Abbreviations: BOADICEA, the Breast and Ovarian Analysis of Disease Incidence and Carrier Estimation Algorithm; NCCN, the National Comprehensive Cancer Network guideline; PRS, Polygenic Risk Score; PV, Pathogenic Variant.

Table S8: Breast cancer lifetime risk category change based on the NICE guideline

Group	BOADICEA Lifetime risk		No gene-test result		Non-PV carriers		*CHEK2* PV carriers[a]		*ATM* PV carriers[a]		*PALB2* PV carriers	
	Without PRS_{313}	Including PRS_{313}	N	% change	N	% change	N	% change	N	% change	N	% change
Cases	**<17%**	**<17%**	478	38.5	699	37.1	1	0.0	NA		NA	
		>17%	299		413		0					
		17-30%	332	34.3	799	31.5	34	48.5	0	100.0	NA	
	17-30%	**<17%**	68		203		1		0			
		>30%	105		164		31		5			
	>30%	**>30%**	42	14.3	65	28.6	93	7.0	32	5.9	10	0.0
		<30%	67		26		7		2		0	
		Overall change	**36.0**		**34.0**		**23.4**		**17.9**		**0.0**	
Controls	**<17%**	**<17%**	783	12.0	2,289	9.8	NA		NA		NA	
		>17%	107		248							
		17-30%	NA		NA		20	35.5	5	44.4	NA	
	17-30%	**<17%**					5		0			
		>30%					6		4			
	>30%	**>30%**	NA		NA		NA		NA		7	0.0
		<30%									0	
		Overall change	**12.0**		**9.8**		**35.5**		**44.4**		**0.0**	

[a]Two cases with both a pathogenic variant in *CHEK2* and *ATM* were excluded.

In total, 1,331 cases and 890 controls were included without a gene-test result; 2,369 cases and 2,537 controls in the non-PV carrier group; 167 cases and 31 controls in the CHEK2 PV carrier group; 39 cases and 9 controls in the ATM carrier group; 10 cases and 7 controls in the PALB2 PV carrier group.

Abbreviations: BOADICEA, the Breast and Ovarian Analysis of Disease Incidence and Carrier Estimation Algorithm; NICE, the National Institute for Health and Care Excellence guideline; PRS, Polygenic Risk Score; PV, Pathogenic Variant.

Table S9: Breast cancer lifetime risk by age of breast cancer diagnosis for cases based on the Dutch IKNL guideline

Group		<40 years		40-50 years		≥50 years	
	BOADICEA LTR	Without PRS_{313}	Including PRS_{313}	Without PRS_{313}	Including PRS_{313}	Without PRS_{313}	Including PRS_{313}
No gene-test result	**<20%**	403 (87%)	305 (66%)	377 (74%)	257 (50%)	222 (62%)	172 (48%)
	20-30%	58 (13%)	127 (27%)	111 (22%)	186(36%)	111 (31%)	122 (34%)
	>30%	1 (0%)	30 (6%)	24 (5%)	69 (13%)	24 (7%)	63 (17%)
Non-PV carriers	**<20%**	475 (81%)	367 (62%)	706 (65%)	557 (52%)	431 (61%)	354 (50%)
	20-30%	96 (16%)	183 (31%)	328 (30%)	395 (37%)	242 (34%)	267 (38%)
	>30%	17 (3%)	38 (6%)	44 (4%)	126 (12%)	30 (4%)	82 (12%)
CHEK2 PV carriers[a]	**<20%**	4 (8%)	3 (6%)	4 (5%)	1 (1%)	2 (4%)	3 (7%)
	20-30%	17 (35%)	12 (24%)	22 (30%)	11 (15%)	18 (40%)	13 (29%)
	>30%	28 (57%)	34 (69%)	47 (46%)	61 (84%)	25 (56%)	29 (64%)
ATM PV carriers[a]	**<20%**	NA	NA	NA	NA	NA	NA
	20-30%	2 (20%)	1 (10%)	2 (12%)	1 (6%)	1 (8%)	0 (0%)
	>30%	8 (80%)	9 (90%)	15 (88%)	16 (94%)	11 (92%)	12 (100%)
PALB2 PV carriers	**<20%**	NA	NA	NA	NA	NA	NA
	20-30%	NA	NA	NA	NA	NA	NA
	>30%	4 (100%)	4 (100%)	5 (100%)	5 (100%)	1 (100%)	1 (100%)

[a]Two cases with both a pathogenic variant in *CHEK2* and *ATM* were excluded.
In total, 1,331 cases were included without a gene-test result; 2,369 cases in the non-PV carrier group; 167 cases in the CHEK2 PV carrier group; 39 cases in the ATM carrier group; 10 cases in the PALB2 PV carrier group.
Abbreviations: BOADICEA, the Breast and Ovarian Analysis of Disease Incidence and Carrier Estimation Algorithm; IKNL, Netherlands Comprehensive Cancer Organisation guideline; LTR, Life Time Risk; PRS, Polygenic Risk Score; PV, Pathogenic Variant.

Supplementary references

1. van Asperen CJ, Brohet RM, Meijers-Heijboer EJ, et al. Cancer risks in BRCA2 families: estimates for sites other than breast and ovary. *Journal of medical genetics*. Sep 2005;42(9):711-9. doi:10.1136/jmg.2004.028829
2. Schmidt MK, Hogervorst F, van Hien R, et al. Age- and Tumor Subtype-Specific Breast Cancer Risk Estimates for CHEK2*1100delC Carriers. *Journal of clinical oncology : official journal of the American Society of Clinical Oncology*. Aug 10 2016;34(23):2750-60. doi:10.1200/jco.2016.66.5844
3. Liu J, Prager-van der Smissen WJ, Schmidt MK, et al. Recurrent HOXB13 mutations in the Dutch population do not associate with increased breast cancer risk. *Sci Rep*. Jul 18 2016;6:30026. doi:10.1038/srep30026
4. Mavaddat N, Michailidou K, Dennis J, et al. Polygenic Risk Scores for Prediction of Breast Cancer and Breast Cancer Subtypes. *American journal of human genetics*. Jan 3 2019;104(1):21-34. doi:10.1016/j.ajhg.2018.11.002
5. Dorling L, Carvalho S, Allen J, et al. Breast Cancer Risk Genes - Association Analysis in More than 113,000 Women. *The New England journal of medicine*. Jan 20 2021;doi:10.1056/NEJMoa1913948
6. Landrum MJ, Lee JM, Benson M, et al. ClinVar: improving access to variant interpretations and supporting evidence. *Nucleic Acids Res*. Jan 4 2018;46(D1):D1062-d1067. doi:10.1093/nar/gkx1153