Implementation of Risk-Stratified Breast Cancer Prevention With a Polygenic Risk Score Test in Clinical Practice

Breast Cancer: Basic and Clinical Research
Volume 17: 1–13

DOI: 10.1177/11782234231205700

Peeter Padrik[1,2], Mikk Puustusmaa[1], Neeme Tõnisson[1,3,4], Berit Kolk[1], Regina Saar[1], Anna Padrik[1] and Tõnis Tasa[1]

[1]OÜ Antegenes, Tartu, Estonia. [2]Clinic of Hematology and Oncology, Tartu University Hospital, Tartu, Estonia. [3]Institute of Genomics, University of Tartu, Tartu, Estonia. [4]Genetics and Personalized Medicine Clinic, Tartu University Hospital, Tartu, Estonia.

ABSTRACT

BACKGROUND: Breast cancer (BC) screening with mammography reduces mortality but considers currently only age as a risk factor. Personalized risk-based screening has been proposed as a more efficient alternative. For that, risk prediction tools are necessary. Genome-wide association studies have identified numerous genetic variants (single-nucleotide polymorphisms [SNPs]) associated with BC. The effects of SNPs are combined into a polygenic risk score (PRS) as a risk prediction tool.

OBJECTIVES: We aimed to develop a clinical-grade PRS test suitable for BC risk-stratified screening with clinical recommendations and implementation in clinical practice.

DESIGN AND METHODS: In the first phase of our study, we gathered previously published PRS models for predicting BC risk from the literature and validated them using the Estonian Biobank and UK Biobank data sets. We selected the best performing model based on prevalent data and independently validated it in both incident data sets. We then conducted absolute risk simulations, developed risk-based recommendations, and implemented the PRS test in clinical practice. In the second phase, we carried out a retrospective analysis of the PRS test's performance results in clinical practice.

RESULTS: The best performing PRS included 2803 SNPs. The C-index of the Cox regression model associating BC status with PRS was 0.656 (SE = 0.05) with a hazard ratio of 1.66. The PRS can stratify individuals with more than a 3-fold risk increase. A total of 2637 BC PRS tests have been performed for women between the ages 30 and 83. Results in clinical use overlap well with expected PRS performance with 5.7% of women with more than 2-fold and 1.4% with more than 3-fold higher risk than the population average.

CONCLUSION: The PRS test separates different BC risk levels and is feasible to implement in clinical practice.

KEYWORDS: breast cancer, polygenic risk score, screening, personalized prevention, genetic predisposition

RECEIVED: November 13, 2022. **ACCEPTED:** September 18, 2023.

TYPE: Original Research Article

FUNDING: The authors disclosed receipt of the following financial support for the research, authorship, and/or publication of this article: This research has been conducted using the source data from UK Biobank under Application Reference #53602. This activity has received funding from the European Institute of Innovation and Technology (EIT) via EIT Health Digital Sandbox Program, Project #2019-DS1001-6541, and Innovation Activity BRIGHT, project #220720. This body of the European Union receives support from the European Union's Horizon 2020 research and innovation program. Neeme Tõnisson has been supported by a PRG555 grant from Estonian Research Council.

COMPETING INTERESTS: The authors declared the following potential conflicts of interest with respect to the research, authorship, and/or publication of this article: BC PRS test AnteBC has been developed by health technology company OÜ Antegenes. AnteBC is registered as a medical device (IVD) in Estonian Medical Devices Database (EMDDB code: 14726). Peeter Padrik, Tõnis Tasa, Berit Kolk, and Neeme Tõnisson have ownership in OÜ Antegenes.

CORRESPONDING AUTHOR: Peeter Padrik, OÜ Antegenes, Raatuse 21, Tartu 50603, Estonia. Email: peeter.padrik@antegenes.com

Introduction

Breast cancer (BC) is the leading cause of cancer deaths in women. Every year adds 2 million new diagnoses and more than 600 000 deaths.[1] Breast cancer screening with mammography reduces BC mortality risk to 20% to 40%.[2-4] Current BC screening guidelines are mostly based on age only and do not support regular screening of women below the age of 50. In most European countries, women aged 50 to 69 years are invited to BC screening at 2-year intervals.[5,6] Such an approach does not account for the wide variation in individual women's BC risks and disregards younger women with a higher risk, but also women over age 50 with higher risk levels who could benefit from personalized screening. Risk-based screening, in which individualized risk assessment is used to inform screening practices, has been proposed as an alternative to age-based screening.[7,8]

Around 30% of the total BC risk has been shown as hereditary.[9] Genetic factors include rare pathogenic variants (PV) in high- and moderate-risk cancer predisposition genes (*BRCA1*, *BRCA2*, etc), having effects large enough to warrant monogenic testing.[10-12] However, only a fraction (5%-10%) of BC cases are caused by these rare PVs.[13] A considerable part of BC risk variation is explained by variants outside these high-risk genes in the form of BC-associated common single-nucleotide polymorphisms (SNPs), identified by genome-wide association studies (GWAS).[14,15] A polygenic risk score (PRS) is the combined effect of individual BC susceptibility SNPs. Although individual associated SNPs may confer only modest disease risk, the combined effect of all known associated SNPs on risk can be substantial. Breast cancer PRSs identify differences in genetic risks and provide a straightforward basis for

designing personalized screening programs by accounting for individual genetic susceptibility.[16] Currently, PRSs have not yet been implemented in routine BC screening, but simulations have suggested that risk profile informed preventive activities could provide cost savings and health benefits.[17,18] High-risk estimation could be also the indication for the use of hormonal chemoprevention.[19]

There is so far no consensus clinical model for the systematic implementation of PRS in BC personalized screening.[20] The current report describes the development of a PRS test as a clinical tool for BC risk-stratified screening, a model set of recommendations for the clinical implementation and the first results of application in real-life clinical practice.

Methods

We conducted a 2-phase investigation. In the first phase, we retrospectively validated the performance of PRSs using both prevalent and incident data sets from 2 genetic biobanks, with the aim of finding the best performing model for a PRS test. In the second phase, we performed a retrospective analysis of the results obtained during the routine clinical implementation of the test. The development of the PRS test as a clinical tool for the characterization of individual polygenic BC risk is described in detail in Appendix 1.

Study cohorts for the PRS development

Breast cancer data sets with genotyped data were acquired from 2 population biobanks: the Estonian Biobank of the Estonian Genome Center at the University of Tartu (EstBB) and the UK Biobank (UKBB). Quality-controlled samples were divided into prevalent and incident data sets. In the EstBB cohort, we retained a total of 32 548 quality-controlled female samples. The prevalent data set contained 315 cases of BC that were diagnosed before biobank recruitment and 1602 controls. The incident data set contained 365 cases of BC that were diagnosed after biobank recruitment and 30 266 controls. The UKBB data set contained 249 062 samples that passed the quality controls. In the UKBB, we identified 8637 prevalent cases and 6825 incident cases that were complemented with 44 952 controls and 188 648 controls, respectively. Prevalent data sets were used for identifying the best candidate model and the incident data sets were used to obtain an independent PRS effect estimate on BC status.

Selection and analysis of candidate PRS models

The literature search for PRS models in the public domains was performed with Google Scholar and PubMed web search engines. A list of articles using the search ["Polygenic risk score" or "genetic risk score" and "breast cancer"] were manually checked for the inclusion criteria.

We evaluated the relationship between BC status and standardized PRS in the 2 prevalent data sets with a logistic regression model to estimate the logistic regression–based odds ratio per 1 standard deviation of PRS (OR_{sd}), its p value, model Akaike information criterion (AIC), and area under the receiver operating characteristic (ROC) curve (AUC). We also pruned the PRS from multi-allelic, non-autosomal, non-retrievable variants based on bioinformatics re-analysis with Illumina GSA-24v1 genotypes and non-overlapping variants between EstBB and UKBB data.

We selected the candidate model with the highest AUC to independently assess risk stratification in the incident data sets.

The main aim of the analyses in the incident data sets was to derive a primary risk stratification estimate, hazard ratio per 1 unit of standardized PRS (HR_{sd}), using a right-censored and left-truncated Cox regression survival model. We also assessed the goodness-of-fit of the survival model using Harrell's C-index and the likelihood ratio test.

Furthermore, we evaluated the concordance between theoretical hazard ratio estimates derived with the continuous per unit PRS (HR_{sd}) estimate and the hazard ratio estimates inferred empirically from data.

For individual BC absolute risk 10-year calculations, we used the risk model developed by Pal Choudhury et al.[21] This absolute risk model allows disease background data from any country. In the current analysis, we used Estonian background information. Theoretical proportions of individuals belonging to risk groups were derived by extracting relative risk estimates for PRS percentiles 1 to 100 from the Choudhury et al model. Conformance of counts of individuals belonging to risk groups to theoretical values was performed using 2-tailed exact binomial tests.

Detailed methods and data sources for that are characterized in Appendix 1.

Development of the clinical implementation model and clinical recommendations

We evaluated PRS risk stratification in the Estonian BC screening context and simulated the extent of risk separation in the Estonian population. Women in Estonia currently start BC screening at age 50. Our analysis first established the 10-year risk of a 50-year-old woman with a population average of PRS ("average female") using the model by Pal Choudhury et al[21]: the reference for the level of risk that initiates population-level screening. Here, we assessed the differences in ages where individuals in various PRS risk percentiles attain 1-fold to 3-fold risk increases compared with the 10-year risk of an average woman.

Based on these analyses, we developed recommendations for a BC screening attendance program based on prescreening PRS testing. This approach uses both relative risks, a fold difference of 10-year risks compared with a genetically average woman of the same age, and her absolute 10-year risk.

The technical documentation and the whole testing pipeline were created, and the BC PRS test was registered as a

Table 1. Comparison metrics of breast cancer PRS models based on the prevalent Estonian Genome Center and UK Biobank data sets.

Variants included in actually tested models		257	73[15]	78[24]	2803[23]
Variants in original PRS		313[23]	77	78	3820
Estonian Genome Center	AUC (SE)	0.604 (0.039)	0.591 (0.004)	0.573 (0.039)	0.615 (0.039)
	OR_{sd} (95% CI)	1.43 (1.27-1.61)	1.37 (1.22-1.54)	1.27 (1.13-1.43)	1.47 (1.31-1.65)
	AIC	1681.2	1689.6	1701.7	1674.1
	McFadden/Nagelkerke Pseudo-R^2	2.1%/3.1%	1.9%/2.0%	0.8%/1.3%	2.5%/3.7%
UK Biobank	AUC (SE)	0.625 (0.0073)	0.607 (0.0075)	0.584 (0.0073)	0.632 (0.0072)
	OR_{sd} (95% CI)	1.55 (1.51-1.59)	1.48 (1.45-1.52)	1.36 (1.33-1.39)	1.62 (1.58-1.66)
	AIC	45900	46273	46705	45695
	McFadden/Nagelkerke Pseudo-R^2	3.0%/4.5%	2.2%/3.3%	1.3%/2.0%	3.5%/5.1%

Abbreviations: PRS, polygenic risk score; AUC, area under the receiver operating characteristic curve; OR, odds ratio; CI, confidence interval; AIC, Akaike information criterion.

medical device (IVD) in Estonian Medical Devices Database (EMDDB code: 14726). For clinical use, a laboratory test report (sample in Appendix 2) and a more detailed report for written post-test counseling were created. The test was implemented on the clinical-grade level as a medical device–based healthcare service by OÜ Antegenes or by partner health care institutions in Estonia. The test is used as home-based testing or at a clinical site taken test with a buccal swab for DNA extraction or using existing genotyped data in the Estonian Biobank. The standard procedure includes the use of Illumina Global Screening Array-24 (GSA) v3.0 chip and Illumina iSCAN sequencer for genotyping. This workflow genotypes ~762 000 markers on the GSA chip by Illumina's Infinium HTS (high-throughput screening) protocol (Illumina Inc, http://www.illumina.com; Document 15045738 v04). The test can also use information from other microarrays and sequencing approaches that output DNA data broadly covering the human genome. The PRS test performs the risk assessment based on imputed genotype data. Quality-controlled markers resulting from genotyping are imputed using a 1000G panel with reference to the human genome GRCh37. The HRC and TOPMed panels were not adopted due to the requirement of using external imputation servers.

Each test was preceded by informed consent. In addition to written information, pretest and posttest oral counseling were offered if needed. All women received PRS test reports directly via the web portal, but the results are also transmitted to the national health information system, where they are also available to other health care providers. According to the test results and recommendations, women are referred to mammography screening institutions or breast clinics for further personalized screening. For additional PV testing recommendations, we use a questionnaire about family cancer history with recommended indications for PV testing.[22] Current official BC screening in Estonia is for the age group 50 to 69 years; for the analysis, we divide women into 2 groups: age 30 to 49 (prescreening age) and 50 to 83 (screening age and older) years.

Results

Selection of the PRS model for the clinical PRS test

Altogether, we chose 4 models from 3 different articles to be evaluated: PRS313,[23] PRS77,[15] PRS78,[24] and PRS3820[23] (original numbers of variants), Table 1. The best performing model was selected based on AUC, OR_{sd}, AIC, and pseudo-R^2 metrics in both EstBB and UKBB data. The PRS 2803 model (that was based on Mavaddat et al PRS3820 model)[23] performed the best (Table 1).

Development of a set of clinical recommendations

To apply a PRS in clinical practice, testing must be complemented by clinical recommendations for preventive activities at different levels of risk. Based on developed PRS induced risk differences, we developed personalized recommendations that are based on relative risks compared with an individual of the same nationality, age, and sex, and the estimated absolute risks. Recommendations presented in Table 2 are based on the age when an individual attains the risk level of genetically average 50-year-old women, as the average risk level at the age of 50 is a generally accepted recommendation to start BC screening, and also using the analogy with clinical recommendations for moderate-risk PV carriers.[22]

Table 2. Recommendations for personalized screening based on different PRS risk levels.

1. Relative risk is less than the population average
   a. Standard mammography screening starting from age 50[a]
2. Relative risk increases up to 2-fold
   a. Recommended mammography screening initiation (2-year interval) at the age of attaining a 10-year risk equivalent to that of a genetically average 50-year-old[a]
3. Relative risk increases 2-fold to 3-fold
   a. Recommend 2a
   b. Recommend mammography screening initiation with 1-year interval starting from an age where 10-year risk attains 2-fold of a genetically average 50-year-old's risk[a]
   c. Besides 3.a and 3.b, we recommend discussing the usage of BC risk-reducing hormonal chemoprevention (tamoxifen, aromatase inhibitors) with a specialist
4. Relative risk increases more than 3-fold
   a. Recommend 3a and 3b
   b. At the age of attaining more than 3-fold of a genetically average 50-year-old's risk then recommend magnetic resonance imaging every 1-2 years[a]
   c. Besides, to 4.a and 4.b, we recommend discussing the usage of BC risk–reducing hormonal chemoprevention (tamoxifen, aromatase inhibitors) with a specialist

Abbreviations: PRS, polygenic risk score; BC, breast cancer.
[a]If the recommended age is below the individual's current age, then we recommend the current age.

Table 3. Women's distribution according to breast cancer PRS risk levels in clinical practice in Estonia and comparison with theoretically calculated risk distribution percentages (CI using 2-sided exact binomial test).

			ALL WOMEN AGE 30-83		WOMEN AGED 30-49		WOMEN AGED 50-83	
N			2637		1881		756	
Mean age			44.7		39.2		58.3	
RISK GROUPS	RISK LEVELS		N	% (95% CI)	N	% (95% CI)	N	% (95% CI)
1	Relative risk is less than the population average	Theoretical	1319	50	941	50	378	50
		Actual	1361	51.6 (49.7-53.5)	979	52.0 (49.8-53.5)	382	50.5 (46.9-54.2)
2	Relative risk increases up to 2-fold	Theoretical	1093	42	779	42	313	42
		Actual	1088	41.3 (39.4-43.2)	761	40.5 (38.2-42.7)	327	43.3 (39.7-46.9)
3	Relative risk increases 2-fold to 3-fold	Theoretical	186	7	132	7	53	7
		Actual	150	5.7[a] (4.8-6.6)	110	5.8 (4.8-7.0)	40	5.3 (3.8-7.1)
4	Relative risk increases more than 3-fold	Theoretical	40	1	28	1	11	1
		Actual	38	1.4 (1.0-2.0)	31	1.6 (1.1-2.3)	7	0.9 (0.4-1.9)

Abbreviations: PRS, polygenic risk score; CI, confidence interval.
[a]Two-sided exact binomial test p value < 0.05

Results from the clinical implementation

Between September 1, 2020, and February 28, 2022, 2637 BC PRS tests with clinical recommendations have been performed for women between the ages of 30 and 83 in the Estonian health care setting. Patients' distribution into risk groups and the comparison with theoretical PRS risk distributions are characterized in Table 3. As the PRS distributes individuals on the normal distribution curve of the risk, it is possible to calculate the theoretical numbers of individuals at different risk levels.

In the age group of 30 to 49 years, BC PRS test has been applied to 1881 women. Current testing has detected 318 (16.9%) women, whose risk level is already as high or higher than average at age 50. 141 (7.5%) of women had risk levels 2 times higher than average and who could discuss more frequent mammographies and hormonal chemoprevention. Thirty-one (1.6%) tested women had risk levels 3 times higher than average and who are candidates for screening magnetic resonance imaging (MRI).

In the age group 50 to 83 years, BC PRS test has been applied to 756 women. We detected 77 women (6.2%) who

have 2 times higher PRS risk levels and are candidates for annual screening mammography and for hormonal chemoprevention. Seven women (0.9%) had risk levels 3 times higher than average; they got recommendations for additional MRI screening.

In combined analysis of all tested women, we note a small, but statistically significant (binomial test $p = 0.0075$) underrepresentation (n observed = 150; n expected = 186) of women belonging to risk group 3 (2-3× increase in relative risk). This difference is present, but not statistically significant in both 30 to 49 and 50 to 83 age groups and is not statistically significant after Bonferroni correction for 12 tests. No significant deviations from expectation were observed in other age and risk groups.

Discussion

Risk-based BC screening, in which individualized risk assessment is used to inform screening practices, has been proposed as an alternative to age-based screening.[7,8] For that, risk prediction tools are necessary. Age only is an imperfect marker for BC risk, given that genetic susceptibility, lifestyle factors, and reproductive history can affect a woman's chance of developing BC. As around 30% of the total BC risk has been shown as hereditary,[9] evaluation of genetic predisposition for BC can serve as a tool for risk-stratified screening.

A considerable part of BC genetic variation is explained by SNPs,[14,15] effects of which are summarized into PRS. Breast cancer PRS identifies differences in genetic risks and provides a straightforward basis for designing personalized screening programs by accounting for individual genetic susceptibility.[16] Currently, PRS scores have not yet been implemented in routine BC screening.[20] As a relatively small proportion of the population are carriers of BC PVs, we have aimed to develop an additional BC risk prediction tool in the form of PRS.

In this study, we validated different publicly available PRS models to find the best performing model for predicting the risk of BC. Our best performing model, named PRS2803, was a pruned version of Mavaddat et al 3820 PRS model containing a total of 2803 SNPs out of 3820.[23] Its performance was consistent with the author's results. Our model was used to design a novel absolute risk–based screening strategy. It is based on Estonian screening information and background data to identify the extent of more than 10-fold PRS-based risk differences between the extremes. Our analysis showed that 1% of women would need to join screening by the age of 34 and more than 30% of individuals do not ever attain the risks of a genetically average 50-year-old woman (the age when women conventionally start screening).

For the clinical implementation, we developed recommendations based mostly on 2 aspects—the average BC risk level currently accepted for routine BC public screening and the analogy with recommendations for moderate-risk PV carriers.[22] Using PRS, it is possible to divide the patient's relative risk of developing BC into different levels compared with the average in the given age while accurately assessing the risk of a particular percentile. In Europe, mammography screening in the age group 50 to 69 years at 2-year intervals is currently a recognized standard practice, which reduces BC mortality. Consequently, the "zero point" of the risk level at the beginning of the screening is the average risk level of 50-year-old women. Detecting younger women with similar or higher risk levels already from age 30 to 35 allows implementing similar mortality reduction measures, avoiding same time screening measures for women with lower risk.

As PRS can predict similar BC risk levels to moderate-risk PVs (*ATM, CHEK2*, and others), we used for our recommendations analogy with these for mammography intervals and MRI use.[22] It should be noted that we do not recommend individuals to join public screening programs later than the standard starting time as the potential benefits and losses from decreased intervals have not been separately validated. As the PRS test does not analyze PVs that significantly increase the risk of BC, our application model recommends additional counseling and testing for the PVs according to the widely acknowledged criteria for PV testing.[22] We see that in the future also all women should be tested for PVs, as family history criteria may miss a proportion of actual PV carriers, but currently the main obstacle to that has been the relatively high cost of PV tests.

Individual risk-based approach for BC screening has allowed the development of more equivalence and equitability among women regarding screening. If an average risk at age 50 (or at any other age) is a commonly accepted level for mammography screening, then personalized recommendations to start screening when an individual risk level reaches average risk create equivalence. If some individuals in average risk level are currently offered interventions or screening, then it would be unfair to deny that to others with equivalent or higher risk of disease. It is important to identify individuals who are at high risk but are currently invisible to the system. Also, for all women, the PRS test can also serve as a tool for individual informed "shared decisions" for mammography screening participation.[6]

Results from the clinical application of our approach show that actual results of PRS testing overlap well with expected results (Table 3). The detected underrepresentation of women belonging to risk group 3 can most likely be attributed to sampling error in PRS percentiles 93 to 99. Alternatively, the difference may arise from uncertainties in population-specific mean and standard deviation estimates used for converting the raw PRS values to z-scores. We are continuing to monitor this observation as more data become available.

Several combined risk prediction models incorporate traditional risk factors such as demographics, reproductive history, menopausal status, family history, previous biopsies, mammographic density, carrier status of PV, and PRS.[25-28] The practical

routine application of such compounded models in screening is complicated due to the nonavailability and quality of data. In practical settings, the data collection difficulties need to be weighted with expected gains. The feasibility, clinical utility, costs, and cost-effectiveness of risk-based programs using a comprehensive model versus a model with only PRS and additional PV testing need to be evaluated additionally.[20] Polygenic risk score alone has been shown to predict the risk of BC in European descent individuals more accurately than current clinical models.[29] Van den Broek et al. have assessed the clinical utility of a first-degree BC family history and PRS to inform screening decisions among women aged 30 to 50 years.[30] Results suggested that BC family history and PRS could guide screening decisions before age 50 years among women at increased risk of BC with the potential to prevent more BC deaths for identifiable groups of women at high risk due to their BC family history and polygenic risk. Analysis by Wolfson et al. concluded that population-wide programs for BC screening that seek to stratify women by their genetic risk should focus first on PRS, not on more highly penetrant but rarer variants, or family history.[31] The PRS was most predictive for identifying women at high risk, while family history was the weakest.

The weakness of our approach is that this is not yet based on the results of randomized trials. Randomized clinical trials of screening interventions provide the strongest evidence of efficacy, although they have certain limitations—a long time to perform and accordingly uncertainty about the relevance of the original study approach after a long study period due to additional scientific progress. Therefore, simulations and modeling studies can indicate which screening strategies are likely to be optimal in each setting. Such modeling studies are started based on the developed approach described in the current report.

In conclusion, we have used a PRS-based model to develop a novel model for BC screening and implemented that with a questionnaire for additional stronger risk MVP testing in clinical practice for personalized recommendations. Our adapted PRS model identifies individuals at more than 3-fold risk and elucidates large differences in attaining the same level of absolute risk. The genetic risk-based recommendations can be applied prospectively by individuals and by institutions aiming to make screening provisions more efficient. Our approach is easily adaptable to other nationalities by using population background information data of other genetically similar populations. For different ethnicities, additional ethnicity-based validations are necessary. We have implemented current BC PRS in Estonia for women with European ethnic background, as GWAS analyzed for current development were based on ethnicities with a European background. Similarly, the clinical screening recommendations can be adapted to locality-specific screening environments if we can infer the absolute risk of the average woman in that locality.

Conclusions

In the additional validation of BC PRSs in EstBB and UKBB, the model with 2803 SNPs demonstrated improved performance compared with models with a smaller number of SNPs:

- PRS test separates different BC risk levels.
- BC PRS test is feasible to implement in clinical practice for risk-stratified BC prevention.

Declarations

Ethics approval and consent to participate

Estonian Biobank (EstBB): Broad consent signed by the EstBB by participants at joining the biobank permits for continuous updates of epidemiologic and health data through periodical linking to Estonian electronic health databases and hospital information systems. The study was conducted in accordance with good ethical standards, and was approved by the Ethics Committee of the University of Tartu (protocol nr 234/T-12). UK Biobank: The UK Biobank study was approved by the North West Multi-Centre Research Ethics Committee (UK Biobank reference: 16/NW/0274). All participants provided written informed consent to participate in the UK Biobank study.

Consent for publication

Not applicable

Author contribution(s)

Peeter Padrik: Conceptualization; Data curation; Formal analysis; Funding acquisition; Investigation; Methodology; Project administration; Resources; Supervision; Validation; Writing—original draft; Writing—review & editing.

Mikk Puustusmaa: Conceptualization; Data curation; Formal analysis; Investigation; Methodology; Project administration; Resources; Software; Validation; Writing—original draft; Writing—review & editing.

Neeme Tõnisson: Conceptualization; Formal analysis; Investigation; Methodology; Supervision; Writing—review & editing.

Berit Kolk: Conceptualization; Data curation; Investigation; Methodology; Project administration; Resources; Writing—review & editing.

Regina Saar: Data curation; Formal analysis; Investigation; Methodology; Software; Writing—review & editing.

Anna Padrik: Conceptualization; Data curation; Formal analysis; Investigation; Methodology; Project administration; Writing—review & editing.

Tõnis Tasa: Conceptualization; Data curation; Formal analysis; Funding acquisition; Investigation; Methodology; Project administration; Resources; Software; Validation; Visualization; Writing—original draft; Writing—review & editing.

Acknowledgements

Our appreciation goes to everybody from the EstBB actively involved in and supporting the development and implementation of BC PRS into clinical practice. Also, we would not have been able to perform any analyses without the computation resources provided by the HPC Center of the University of Tartu. We would like to extend our gratitude to Siim Sõber for his invaluable contributions to the statistical analyses and further conceptualization in this scientific manuscript.

Availability of data and materials

Individual-level genotype and phenotype data from Estonian Biobank or UK Biobank cannot be explicitly shared. The UK Biobank Resource was used under Application Reference Number 53602. New users can request access to UK Biobank from https://www.ukbiobank.ac.uk/enable-your-research/apply-for-access. Similarly, Estonian Biobank data are available by request. Researchers interested in Estonian Biobank can request the access here: https://genomics.ut.ee/en/content/estonian-biobank.

ORCID iD

Peeter Padrik https://orcid.org/0000-0002-1915-4806

REFERENCES

1. Ferlay J, Colombet M, Soerjomataram I, et al. Estimating the global cancer incidence and mortality in 2018: GLOBOCAN sources and methods. *Int J Cancer*. 2019;144:1941-1953. doi:10.1002/ijc.31937.
2. Myers ER, Moorman P, Gierisch JM, et al. Benefits and harms of breast cancer screening. *JAMA*. 2015;314:1615-1634. doi:10.1001/jama.2015.13183.
3. Tabár L, Dean PB, Chen TH-H, et al. The incidence of fatal breast cancer measures the increased effectiveness of therapy in women participating in mammography screening. *Cancer*. 2018;125:515-523. doi:10.1002/cncr.31840.
4. Independent UK Panel on Breast Cancer Screening. The benefits and harms of breast cancer screening: an independent review. *Lancet*. 2012;380:1778-1786. doi:10.1016/S0140-6736(12)61611.
5. World Health Organization. *WHO Position Paper on Mammography Screening*. Geneva, Switzerland: World Health Organization; 2014.
6. European guidelines on breast cancer screening and diagnosis, 2021.
7. Schousboe JT, Kerlikowske K, Loh A, et al. Personalizing mammography by breast density and other risk factors for breast cancer: analysis of health benefits and cost-effectiveness. *Ann Intern Med*. 2011;155:10-20. doi:10.7326/0003-4819.
8. Shieh Y, Eklund M, Sawaya GF, Black WC, Kramer BS, Esserman LJ. Population-based screening for cancer: hope and hype. *Nat Rev Clin Oncol*. 2016;13:550-565. doi:10.1038/nrclinonc.2016.50.
9. Mucci LA, Hjelmborg JB, Harris JR, et al. Familial risk and heritability of cancer among twins in Nordic Countries. *JAMA*. 2016;315:68-76. doi:10.1001/jama.2015.17703.
10. Frey JD, Salibian AA, Schnabel FR, Choi M, Karp NS. Non-BRCA1/2 breast cancer susceptibility genes: a new frontier with clinical consequences for plastic surgeons. *Plast Reconstr Surg Glob Open*. 2017;5:e1564. doi:10.1097/gox.0000000000001564.
11. Wunderle M, Olmes G, Nabieva N, et al. Risk, prediction and prevention of hereditary breast cancer—large-scale genomic studies in times of big and smart data. *Geburtshilfe Frauenheilkd*. 2018;78:481-492. doi:10.1055/a-0603.
12. Slavin TP, Maxwell KN, Lilyquist J, et al. The contribution of pathogenic variants in breast cancer susceptibility genes to familial breast cancer risk. *NPJ Breast Cancer*. 2017;3:22. doi:10.1038/s41523-017.
13. Apostolou P, Fostira F. Hereditary breast cancer: the era of new susceptibility genes. *Biomed Res Int*. 2013;2013:747318.
14. Ghoussaini M, Pharoah PD. Polygenic susceptibility to breast cancer: current state-of-the-art. *Future Oncol*. 2009;5:689-701.
15. Mavaddat N, Pharoah PD, Michailidou K, et al. Prediction of breast cancer risk based on profiling with common genetic variants. *J Natl Cancer Inst*. 2015;107:djv036. doi:10.1093/jnci/djv036.
16. Lewis CM, Vassos E. Prospects for using risk scores in polygenic medicine. *Genome Medicine*. 2017;9:96. doi:10.1186/s13073-017.
17. Pashayan N, Morris S, Gilbert FJ, et al. Cost-effectiveness and benefit-to-harm ratio of risk-stratified screening for breast cancer: a life-table model. *JAMA Oncol*. 2018;4:1504-1510.
18. Maas P, Barrdahl M, Joshi AD, et al. Breast cancer risk from modifiable and nonmodifiable risk factors among white women in the United States. *JAMA Oncol*. 2016;2:1295-1302. doi:10.1001/jamaoncol.2016.1025.
19. Owens DK, Davidson KW, Krist AH, et al. Medication use to reduce risk of breast cancer: US preventive services task force recommendation statement. *JAMA*. 2019;322:857-867. doi:10.1001/jama.2019.11885.
20. Pashayan N, Antoniou AC, Ivanus U, et al. Personalized early detection and prevention of breast cancer: ENVISION consensus statement. *Nat Rev Clin Oncol*. 2020;17:687-705. doi:10.1038/s41571-020.
21. Pal Choudhury P, Maas P, Wilcox A, et al. iCARE: an R package to build, validate and apply absolute risk models. *PLoS ONE*. 2020;15:e0228198.
22. NCCN clinical practice guidelines in oncology: genetic/familial high-risk assessment: breast, ovarian, and pancreatic, 2021. https://www.nccn.org/professionals/physician_gls/pdf/genetics_bop.pdf.
23. Mavaddat N, Michailidou K, Dennis J, et al. Polygenic risk scores for prediction of breast cancer and breast cancer subtypes. *Am J Hum Genet*. 2019;104:21-34. doi:10.1016/j.ajhg.2018.11.002.
24. Fritsche LG, Gruber SB, Wu Z, et al. Association of polygenic risk scores for multiple cancers in a phenome-wide study: results from the Michigan Genomics Initiative. *Am J Hum Genet*. 2018;102:1048-1061. doi:10.1016/j.ajhg.2018.04.001.
25. Costantino JP, Gail MH, Pee D, et al. Validation studies for models projecting the risk of invasive and total breast cancer incidence. *J Natl Cancer Inst*. 1999;91:1541-1548. doi:10.1093/jnci/91.18.1541.
26. Gail MH, Brinton LA, Byar DP, et al. Projecting individualized probabilities of developing breast cancer for white females who are being examined annually. *J Natl Cancer Inst*. 1989;81:1879-1886. doi:10.1093/jnci/81.24.1879.
27. Tyrer J, Duffy SW, Cuzick J. A breast cancer prediction model incorporating familial and personal risk factors. *Stat Med*. 2004;23:1111-1130. doi:10.1002/sim.1668.
28. Lee A, Mavaddat N, Wilcox AN, et al. BOADICEA: a comprehensive breast cancer risk prediction model incorporating genetic and nongenetic risk factors. *Genet Med*. 2019;21:1708-1718. doi:10.1038/s41436-018.
29. Martin AR, Kanai M, Kamatani Y, et al. Clinical use of current polygenic risk scores may exacerbate health disparities. *Nat Genet*. 2019;51:584-591. doi:10.1038/s41588-019.
30. van den Broek JJ, Schechter CB, van Ravesteyn NT, et al. Personalizing breast cancer screening based on polygenic risk and family history. *J Natl Cancer Inst*. 2021;113:434-442. doi:10.1093/jnci/djaa127.
31. Wolfson M, Gribble S, Pashayan N, et al. Potential of polygenic risk scores for improving population estimates of women's breast cancer genetic risks. *Genet Med*. 2021;23:2114-2121. doi:10.1038/s41436-021.
32. Loh PR, Palamara PF, Price AL. Fast and accurate long-range phasing in a UK Biobank cohort. *Nat Genet*. 2016;48:811-816.
33. Browning BL, Zhou Y, Browning SR. A one-penny imputed genome from next-generation reference panels. *Am J Hum Genet*. 2018;103:338-348.
34. Mitt M, Kals M, Pärn K, et al. Improved imputation accuracy of rare and low-frequency variants using population-specific high-coverage WGS-based imputation reference panel. *Eur J Hum Genet*. 2017;25:869-876.
35. Genotyping and quality control of UK Biobank, a large-scale, extensively phenotyped prospective resource, 2019. https://biobank.ctsu.ox.ac.uk/crystal/crystal/docs/genotyping_qc.pdf.
36. Estonian Health Statistics and Health Research Database, 2020. https://statistika.tai.ee/pxweb/et/Andmebaas/Andmebaas__02Haigestumus.
37. World Health Organization. Global Health Observatory indicator views, 2018. https://apps.who.int/gho/data/node.imr.LIFE_0000000029?lang=en.

Appendix 1

Methodology and results of the polygenic risk score development

Biobank participant data. Breast cancer (BC) data sets were acquired from 2 population biobanks: the Estonian Biobank of the Estonian Genome Center at the University of Tartu (EstBB) and the United Kingdom Biobank (UKBB). Quality-controlled

samples were divided into prevalent and incident data sets. The prevalent data set included BC cases diagnosed before Biobank recruitment with 5 times as many controls without the diagnosis. Incident data included cases diagnosed in any of the linked databases after recruitment to the Biobank and all controls not included in the prevalent data set. Prevalent data sets were used for identifying the best candidate model and the incident data sets were used to obtain an independent PRS effect estimate on BC status.

Participant data in the Estonian Biobank. Breast cancer cases and controls in retrospective data of EstBB were defined by BC ICD-10 (*International Classification of Diseases, Tenth Revision*) code (C50) status derived from questionnaires filled at recruitment of the gene donors and from linked data from Estonian Cancer Registry (data until 2013), National Health Insurance Fund (data until the end of 2018), and Causes of Death Registry (data until 2017 May). All EstBB samples were genotyped in the Core Genotyping Lab of the Institute of Genomics, University of Tartu, using Illumina GSAMD-24v1-0 arrays. Individuals were excluded if their call rate was <95% or sex defined based on X chromosome heterozygosity did not match the declared sex. Variants were filtered by call rate <95% on the whole EstBB data set, Hardy-Weinberg equilibrium (HWE) test p value < 1e-4 (autosomal variants only), and minor allele frequency < 1%. Variant positions were updated to b37 and all variants were changed to the TOP strand (https://www.well.ox.ac.uk/~wrayner/strand/). Phasing was done using Eagle (v 2.3) software[32] and imputation with Beagle (v 28Sep18.793)[33] using the population-specific imputation reference of 2297 WGS samples.[34]

Participant data in the UK Biobank. This study used genotypes from the UKBB cohort (version v3, obtained 07.11.2019) and made available to Antegenes under application reference number 53602. The data were collected and genotyped using either the UK BiLEVE or Affymetrix UK Biobank Axiom Array. Breast cancer cases in the UKBB cohort were retrieved by the status of ICD-10 code C50. We additionally included cases with self-reported UKBB code "1022."

Quality control steps and, in detail, methods applied in imputation data preparation have been described by the UKBB.[35] We applied additional quality controls on autosomal chromosomes. First, we removed all variants with allele frequencies outside 0.1% and 99.9%, genotyping call rate < 0.1, imputation (INFO) score < 0.4, and Hardy-Weinberg equilibrium test p value < 1E-6. Sample quality control filters were based on several predefined UKBB filters. We removed samples with excessive heterozygosity, individuals with sex chromosome aneuploidy, and excess relatives (>10). In addition, we only kept individuals for whom the submitted gender matched the inferred gender, and the genotyping missingness rate was below 5%.

In the EstBB cohort, we retained a total of 32 548 quality-controlled female samples. All samples were divided into prevalent and incident data sets. The prevalent data set contained 315 cases of BC that were diagnosed before Biobank recruitment and 1602 controls. The incident data set contained 365 cases of BC that were diagnosed after Biobank recruitment and 30 266 controls. The UKBB data set contained 249 062 samples that passed the quality controls. In the UKBB, we identified 8637 prevalent cases and 6825 incident cases that were complemented with 44 952 controls and 188 648 controls, respectively.

Model selection from candidate polygenic risk score models. We searched the literature for polygenic risk score (PRS) models in the public domain. The requirements for inclusion in the candidate set were the availability of the chromosomal location, reference and alternative allele, minor allele frequency, and an estimator for the effect size either as odds ratio (OR) or its logarithm (log-OR) specified for each genetic variant. In cases of iterative model developments on the same underlying base data, we retained chronologically newer ones. The search was performed with Google Scholar and PubMed web search engines by working through a list of articles using the search ["Polygenic risk score" or "genetic risk score" and "breast cancer"], and then manually checking the results for the inclusion criteria.

Altogether, we chose 4 models from 3 different articles to be evaluated: PRS313,[23] PRS77,[15] PRS78,[24] and PRS3820.[23] Normality assumption of the standardized PRS was not violated with any tested models (Shapiro-Wilk test p values in EstBB data PRS313 = 0.45, PRS77 = 0.40, PRS78 = 0.28, PRS3820 = 0.46). The best performing model was selected based on AUC, OR_{sd}, AIC, and pseudo-R^2 metrics in both EstBB and UKBB data. The PRS2803 model that was based on Mavaddat et al[23] performed the best (Table 1 in the main text). The corresponding AUC under the ROC curve (Figure A1) for the association between the PRS and BC diagnosis was 0.615 (SE = 0.039) in EstBB and 0.632 (SE = 0.0072) in UKBB.

We additionally pruned the PRS from multi-allelic, non-autosomal, non-retrievable variants based on bioinformatics re-analysis with Illumina GSA-24v1 genotypes and non-overlapping variants between EstBB and UKBB data.

Polygenic risk scores were calculated as $PRS = \sum_{i}^{m} \beta_i \left(\sum_{j=0}^{2} \omega_{ij} xj \right)$, where ω_{ij} is the probability of observing genotype j, *where* $j \in \{0,1,2\}$ for the ith SNP; m is the number of SNPs, and β_i is the effect size of the ith SNP estimated in the PRS. The mean and standard deviation of PRS in the cohort were extracted to standardize individual risk scores to Gaussian. We tested the assumption of normality with the mean of 1000 Shapiro-Wilk test replications on a random subsample of 1000 standardized PRS values.

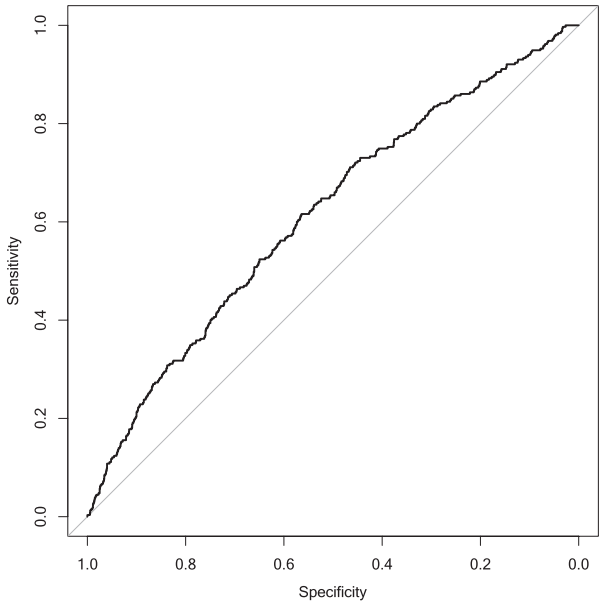

Figure A1. Receiver operating characteristic plot of breast cancer cases and controls in prevalent Estonian Genome Center data set.

Next, we evaluated the relationship between BC status and standardized PRS in the 2 prevalent data sets with a logistic regression model to estimate the logistic regression–based odds ratio per 1 standard deviation of PRS (OR_{sd}), its p value, model Akaike information criterion (AIC), and area under the ROC curve (AUC). The logistic regression model was compared with the null model using the likelihood ratio test and to estimate the Nagelkerke and McFadden pseudo-R^2. We selected the candidate model with the highest AUC to independently assess risk stratification in the incident data sets.

Independent performance evaluation of a PRS modelThe main aim of the analyses in the incident data sets was to derive a primary risk stratification estimate, hazard ratio per 1 unit of standardized PRS (HR_{sd}), using a right-censored and left-truncated Cox regression survival model. The start of time interval was defined as the age of recruitment; follow-up time was set as the time of diagnosis for cases and at the time of last health data linkage for controls. Scaled PRS was used as the only independent variable of BC diagnosis status, and 95% confidence intervals were created using the standard error of the log-hazard ratio. We also assessed the goodness-of-fit of the survival model using Harrell C-index and the likelihood ratio test.

Furthermore, we evaluated the concordance between theoretical hazard ratio estimates derived with the continuous per unit PRS (HR_{sd}) estimate and the hazard ratio estimates inferred empirically from data (Figure A2). For this, we binned the individuals by PRS to deciles and estimated the empiric hazard ratio of BC directly between those classified in each decile and those belonging to the remaining deciles. Theoretically estimated hazard ratio estimates assume a multiplicative effect of the mean in a PRS bin on the unit-based hazard ratio. This relationship between HR_{sd} and the expected mean in the truncated Gaussian PRS distribution is expressed as $HR_{sd}^{\Phi^{-1}(a,b)}$, where the exponent is the mean of a truncated Gaussian distribution between 2 percentiles a and b (bounded between 0 and 1, $a < b$), and $\Phi^{-1}(a,b) = (f(Q(b)) - f(Q(a)))/(a-b)$, where $Q(b)$ is the Gaussian quantile function on a percentile b and $f(Q)$ is the Gaussian probability density function value at a quantile function value. We compared the 2 approaches by using the Spearman correlation coefficient and the proportion of distribution-based $HR_{sd}^{\Phi^{-1}(a,b)}$ estimates in empirical confidence intervals.

Next, we evaluated the performance of the best performing PRS2803 model in the independent incident data sets with the main aim of estimating the hazard ratio per unit of PRS. Table 2 presents the performance estimation metrics. The hazard ratio per 1 unit of standard deviation (HR_{sd}) of model PRS2803 was *1.66* with standard error (log [HR]) = 0.05) in the incident EstBB data set. The concordance index (C-index) of the survival model testing the relationship between PRS and BC diagnosis status in the incident EstBB data set was 0.656 (SE = 0.015) and slightly lower in the UK Biobank.

Hazard ratio estimates in deciles of PRS are visualized in Figure A2. In panel A with EstBB data, the lowest 10% bin includes 3073 controls and 20 cases, whereas in the top 10%, we observed 2946 controls and 74 cases. Respectively, in the UKBB data set (Figure A2(C)) we observed 303 cases and we observed good calibration of PRS-derived hazard ratios to expected values. In the EstBB incident data (Figure A2(B)), a slight overestimation of hazard ratio may be present in the average risk bracket. In the UKBB data set (Figure A2(D)), the hazard ratios are very well aligned with theoretical values (Pearson correlation coefficient = 0.994). The match is particularly exact in above average risk individuals, whereas in the low risk individuals a slight deviation toward lower than expected hazard ratios is noticeable due to the tight confidence intervals in this larger data set.

Absolute risk estimation. Individual τ -year (eg, 10-year) absolute risk calculations are based on the risk model developed by Pal Choudhury et al.[21] Individual absolute risks are estimated for currently a-year-old individuals in the presence of known risk factors (Z) and their relative log hazard ratio parameters (β), and 95% uncertainty intervals for the hazard ratio were derived using the standard error and z-statistic 95% quantiles $CI_{HR} = \exp(\beta \pm 1.96 * se[HR_{sd}])$, where $se(HR_{sd})$ is the standard error of the log-hazard ratio estimate. Risk factors have a multiplicative effect on the baseline hazard function. The model specifies the next τ -year absolute risk for a currently a-year old individual as

$$\int_a^{a+\tau} \lambda_0(t)\exp\left(\beta^T Z\right) exp\left(-\int_a^t \left[\lambda_0(u)\exp\left(\beta^T Z\right) + m(u)\right] du\right) dt,$$

where $m(t)$ is the age-specific mortality rate function and $\lambda_0(t)$ *is* the baseline hazard function, $t \geq T$, and T is the time to onset

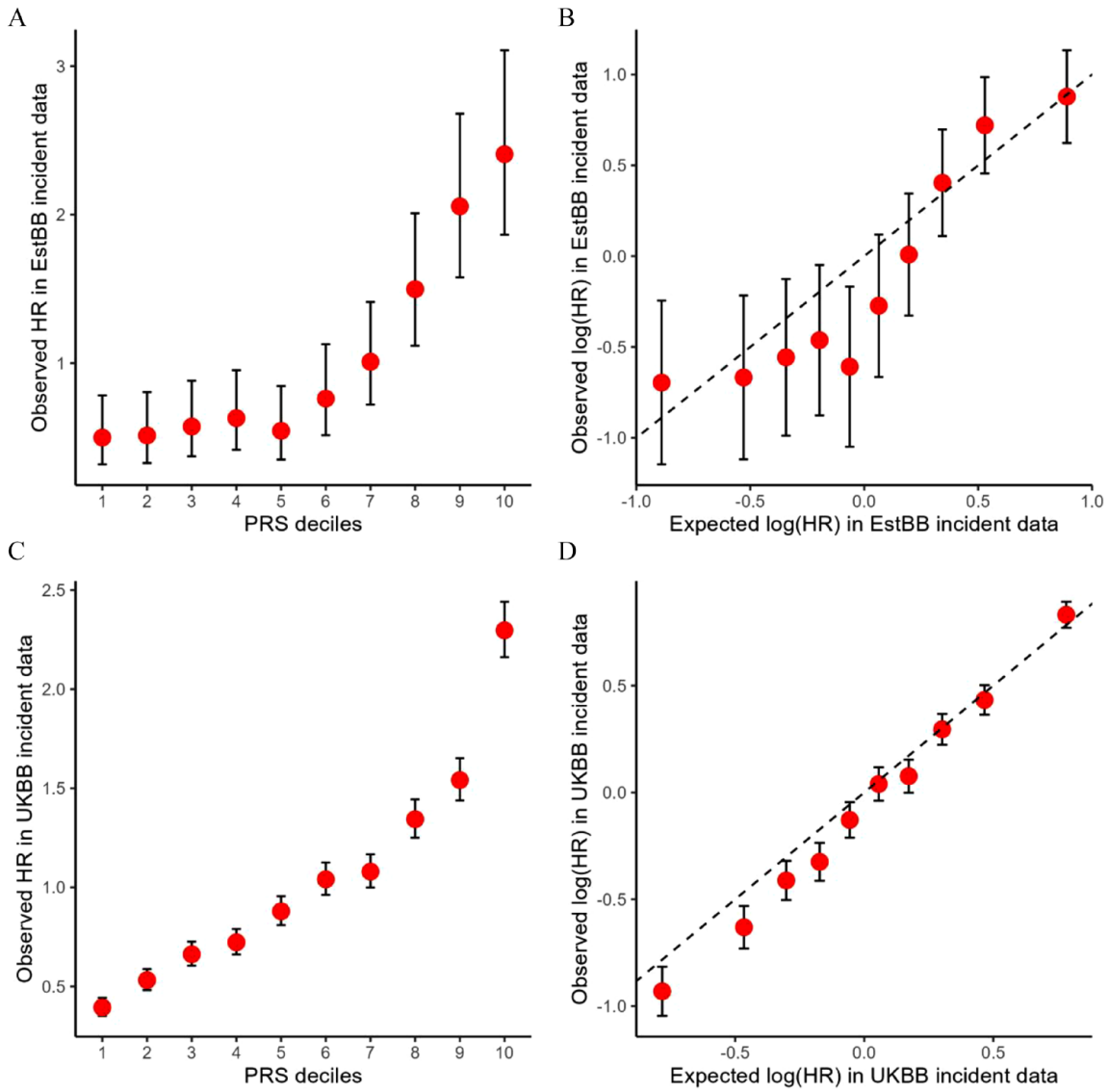

Figure A2. HR estimates between deciles of the PRS in the EstBB and UKBB incident data sets. Red dots and black lines represent empirically estimated HR estimates and corresponding confidence intervals. Black dashes represent the theoretical hazard ratio for the decile bins derived from the HR per unit PRS. (A) EstBB observed HRs in incident data set. (B) Calibration graph of expected and observed HRs in EstBB incident data. (C) UKBB observed HRs in incident data set. (D) Calibration graph of expected and observed hazard ratios in UKBB incident data. EstBB indicates Estonian Biobank; HR, hazard ratio; PRS, polygenic risk scores; UKBB, UK Biobank.

Table A1. Performance metrics of Cox regression model on the disease status and PRS2803-based polygenic risk scores calculated in the incident data sets.

	HR_{SD} (95% CONFIDENCE INTERVAL)	C-INDEX (SE)	−2 × LOG LIKELIHOOD	LIKELIHOOD RATIO TEST *P* VALUE
Estonian Genome Center	1.66 (1.5-1.84)	0.656 (0.015)	95.3	<2e-16
UK Biobank	1.56 (1.53-1.6)	0.625 (0.003)	1351	<2e-16

Abbreviation: HR, hazard ratio.

of the disease. The baseline hazard function is derived from marginal age-specific BC incidence rates (λ_m *([t]*) and the distribution of risk factors Z in the general population ($F(z)$):

$$\lambda_m(t) \approx \int \lambda_0(t) \exp\left(\beta^T Z\right) dF(z).$$

This absolute risk model allows disease background data from any country. In this analysis, we used Estonian background information. We calculated average cumulative risks using data from the National Institute of Health Development of Estonia that provides population average disease rates in age groups of 5-year intervals.

Sample sizes for each age group were acquired from Estonian Health Statistics and Health Research Database for 2013-2016.[36] Next, we assumed constant incidence rates for each year in the 5-year groups. Thus, incidence rates for each age group were calculated as $IR = X_t / N_t$, where X_t is the number of first-time cases at age t and N_t is the total number of women in this age group. Final per-year incidences were averaged over the time range 2013-2016.

Age-and sex-specific mortality data for the year 2016 was retrieved from World Health Organization[37] and competing mortality rates were constructed by subtracting yearly age- and sex-specific disease mortality rates from general mortality rates. Breast cancer mortality estimates were derived from the Global Cancer Observatory.[1]

We applied this model to estimate absolute risks for individuals in the first, 10th, 25th, 50th, 75th, 90th, and 99th PRS quantiles, for example, an individual on the 50th percentile would have a standardized PRS of 0. Confidence intervals for the absolute risk are estimated with the upper and lower confidence intervals of the continuous per unit log-hazard ratio. Similarly, we used the absolute risk model to estimate lifetime risks (between ages 0 and 85) for the individuals in the same risk percentiles.

We used a model by Pal Choudhury et al. to derive individual 10-year risks[21] and specified $F(z)$ as the distribution of PRS estimates in the whole EstBB cohort. The log-hazard ratio (β) is based on the estimate of the log-hazard ratio in the BC16 model of the incident EstBB data set. Age-specific BC incidence and competing mortality rates provided the background for BC incidences in the Estonian population.

In the Estonian population, the absolute risk of developing BC in the next 10 years among 50-year-old women in the first percentile is 0.466% (0.349%-0.616%) and 4.83% in the 99th percentile 4.83% (4.00%-5.77%). At age 70, corresponding risks become 0.59% (0.445%-0.778%) and 6.08% (5.03%-7.30%), respectively. The relative risks between the most extreme percentiles are therefore $>10.3\times$ fold. At the same time, competing risk accounted cumulative risks reach 19.2% by age 85 (16.1%-22.6%) for those in the 99th percentile but remain 2.00% (1.51%-2.64%) for those in the first percentile (Figure A3).

A genetically average 50-year-old woman has a 10-year absolute risk of around 1.51%. PRS2803 model can identify 34-year-old women in the 99th percentile of PRS who have a larger risk than the average risk of 50-year-olds. At the same time, 50-year-old women in the 32nd percentile and lower attain the average risk of 50-year-olds by their 70th birthday. In effect, individual women could be at the risk that currently initiates population-level screening between ages 34 and 70. Similarly, 50-year-old women above the 92nd PRS percentile have a more than 2-fold risk and around 1.3% of women attain a 3-fold risk compared with those at average risk (Figure A4).

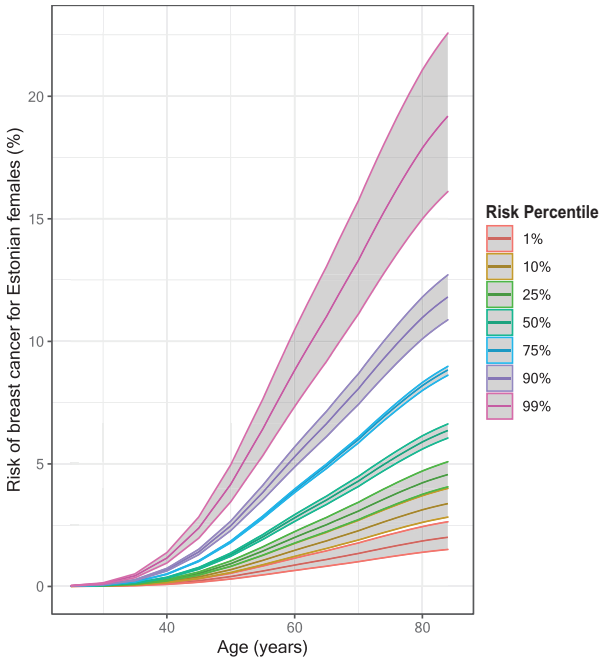

Figure A3. Cumulative risks (%) of breast cancer between ages 20 and 85 in various risk percentiles.

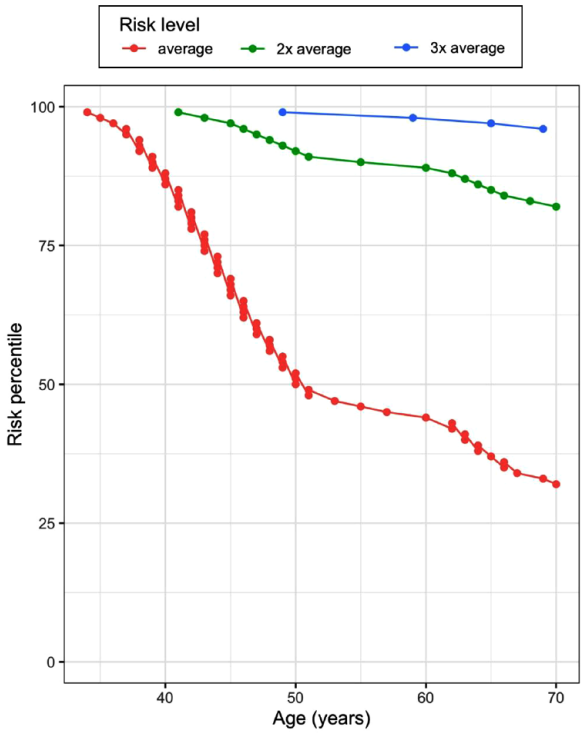

Figure A4. Ages when Estonian women in different risk percentiles attain 1-fold to 3-fold multiples of 10-year risk compared with 50-year-old women with population average polygenic risk score (risk level: "average").

Appendix 2

Sample laboratory report of a breast cancer polygenic risk score test

ANTEBC LABORATORY TEST RESULT
Polygenic Risk Score of Breast Cancer
Last Name: XXX
First Name: XXX
National Identification Number: 41234567890
Age: 40
Analysis code and LOINC code: A-5429 breast cancer polygenic risk score
Clinical Diagnosis: Z01.7 Laboratory Analysis
Analysis/Procedure: Polygenic Risk Score test
Laboratory Test Result: 1.38
Laboratory Analysis Description:
AnteBC test: A-5429 breast cancer polygenic risk score
Sample Material Data
Date of Sample Collection: dd.mm.yyyy
Sample Container Identifier: abcdefg
Sample Material: Mo buccal swab
Analysis Method: Genotyping
Analysis and Interpretation: Antegenes
Laboratory Test Result: The patient's polygenic risk score of breast cancer is 1.38 standard deviation (SD) units.

Interpretation of Laboratory Test Result: The result shows that the breast cancer polygenic risk score is 1.38 standard deviation units higher than the population average placing the patient's risk score among 40-year-old women in the 92nd percentile. Meaning that more than 91% of women have lower and more than 8% of women have higher polygenic risk scores.

The patient's risk of developing BC in the next 10 years is 1.77%. The 10-year average risk of developing breast cancer among 40-year-old women in Estonia is 0.89%. In terms of relative comparison, the current result implies that the risk of developing breast cancer in the next 10 years is 2.00 times higher than the 10-year genetic risk among 40-year-old women on average.

Based on the breast cancer polygenic risk score test results, we recommend,

- Mammography screening every 2 years starting at the age of 40.
- Follow general guidelines to reduce the risk of breast cancer (see our recommendations).
- The individual cancer risk can be high, whether or not there is a family history of cancer. Knowing the personal breast cancer risk is important because further clinical analysis can now be implemented according to the patient's risk. This way we can prevent the disease or detect it as early as possible.

Test results were entered: dd.mm.yyyy
Time of evaluation of the result: dd.mm.yyyy
The results were confirmed: dr. ZZZ ZZZ, D12345. health care professional speciality: E190 Laboratory medicine. Name of health care institution: OÜ Antegenes.

Decision. Based on the breast cancer polygenic risk score test results, we recommend,

- Mammography screening every 2 years starting at the age of 40.
- Follow general guidelines to reduce the risk of breast cancer (see our recommendations).

Notes:
In addition to the polygenic component used by the AnteBC test, there are also other breast cancer risk factors to be considered. Risk assessment could be affected by any previous breast diagnostic tests, occurrences of cancer in close (biological) relatives, or individual health behaviors.

Health behavior

- A body mass index greater than 30 increases the risk of breast cancer by a factor of 1.5 to 2. It is recommended for adults to maintain a body mass index between 18.5 and 24.9 or at least below 30.
- Physical activity is risk-reducing. At least 30 minutes of moderate-to-vigorous intensity physical activity is recommended on most days, a total of 1.5 to 4 hours per week.
- Consuming just one alcoholic drink a day increases breast cancer risk by 5%. Regular alcohol consumption should be avoided to reduce the risk of breast cancer.
- Using hormone replacement therapy (HRT) during menopause increases the risk of breast cancer. The risk of breast cancer increases with the use of estrogen and progestogen combination and estrogen alone. Therefore, the risk-benefit ratio of these drugs should be discussed with your doctor.
- Women who have never given birth and women who give birth to their first child at the age of 35 or older have a slightly higher risk of developing breast cancer.
- Smoking increases the risk of breast cancer.

Body awareness. We recommend you to be aware of your body, including the condition and possible changes in your breasts. If you notice any of the symptoms listed below, we recommend that you seek medical attention. These may indicate the development of breast cancer.

- Abnormal changes in breast shape, size, or color.
- New lump or mass in breast tissue.
- Pain or discomfort in one breast.
- Changes in the surface of the breast skin (looking like an orange peel), skin retraction, "wrinkling," or ulcer.
- Change or retraction of nipple shape or position.
- Bleeding or flushing around the nipple, discharge from the nipple.
- Enlargement of the axillary lymph nodes.

Rare genetic variants. In addition to the polygenic variants considered by AnteBC, there are a few rare variants in the genome that significantly increase the risk of cancer. Such variants are tested in tests for single genes for hereditary tumor risk.

AnteBC test does not analyze rare risk-increasing mutations in single genes. The patient may be indicated for such a test if any of the following criteria are met:

- The patient has a history of breast cancer, ovarian or fallopian tube cancer, or peritoneal cancer;
- The patient's biological relative is known to have a mutation in individual genes for breast and ovarian cancer predisposition (*BRCA1*, *BRCA2*, etc);
- A first- or second-degree biological relative has been diagnosed with breast cancer below the age of 45 years; has been diagnosed with pancreatic cancer, ovarian cancer, metastatic prostate cancer, 2 or more cases of breast cancer in one person; or has a history of breast cancer;
- Biological relatives have had 3 or more tumors associated with hereditary cancer syndromes;
- Ashkenazi Jewish descent.
- In this case, we recommend consulting a Medical Geneticist and testing individual genes for hereditary tumor risk.