Genomic risk prediction for breast cancer in older women

Authors:

Paul Lacaze, PhD¹, Andrew Bakshi, MSc¹, Moeen Riaz, PhD¹, Suzanne G Orchard, PhD¹, Jane Tiller, MGenCouns¹, Johannes T Neumann, MD¹, Prudence R Carr, PhD¹, Amit D Joshi, PhD², Yin Cao, MPH, ScD³, Erica T Warner, MPH², Alisa Manning, PhD², Tú Nguyen-Dumont, PhD⁴, Melissa C. Southey, PhD⁴, Roger L Milne, PhD^{4,5}, Leslie Ford, MD⁶, Robert Sebra, PhD⁷, Eric Schadt, PhD⁷, Lucy Gately, MBBS⁸, Peter Gibbs, MBBS⁸, Bryony A Thompson, PhD⁹, Finlay A Macrae, MBBS⁹, Paul James, MBBS⁹, Ingrid Winship, MBBS⁹, Catriona McLean, MBBS, MD¹⁰, John R Zalcberg, MBBS¹, Robyn L Woods, PhD¹, Andrew T Chan, MD, MPH², Anne M Murray, MD, MSc¹¹ & John J McNeil, MBBS, PhD¹.

Affiliations:

- Department of Epidemiology and Preventive Medicine, School of Public Health and Preventive Medicine, Monash University, Melbourne, Australia.
- Clinical and Translational Epidemiology Unit, Massachusetts General Hospital and Harvard Medical School, Boston, MA, USA; MGH Cancer Center, Boston, MA, USA
- Division of Public Health Sciences, Department of Surgery and Alvin J. Siteman Cancer Center, Washington University School of Medicine, St Louis, MO, USA.
- Precision Medicine, School of Clinical Sciences at Monash Health, Monash University and Clinical Pathology, University of Melbourne, Melbourne, Australia.
- 5. Cancer Epidemiology Division, Cancer Council Victoria, Melbourne, Australia
- 6. National Cancer Institute, Division of Cancer Prevention, Rockville, MD, USA.

- Department of Genetics and Genomic Sciences, Icahn School of Medicine at Mount Sinai, New York, USA
- Personalised Oncology Division, Walter and Eliza Hall Institute Medical Research and Faculty of Medicine University of Melbourne, Australia
- Department of Genomic Medicine; Family Cancer Clinic, Department of Medicine, Department of Pathology, Royal Melbourne Hospital, University of Melbourne, Parkville, VIC, Australia
- 10. Department of Anatomical Pathology, Alfred Hospital, Melbourne, Victoria, Australia.
- 11.Berman Center for Outcomes and Clinical Research, Hennepin Healthcare Research Institute, Hennepin Healthcare, and University of Minnesota, Minneapolis, MN, USA

Corresponding author:

- *Paul Lacaze, PhD. paul.lacaze@monash.edu
- Department of Epidemiology and Preventive Medicine
- School of Public Health and Preventive Medicine
- Monash University, Level 5, The Alfred Centre,
- 99 Commercial Road, Melbourne VIC 3004, Australia.
- Ph. +61 3 9903 0412, Fax: +61 3 9903 0556. ORCID ID: 0000-0002-0902-6798

Word count: 2815

Abstract

Background: Genomic risk prediction models for breast cancer (BC) have been predominantly developed with data from women aged less than 70 years. Prospective studies of women aged 70 years or older have been limited.

Methods: We assessed the effect of a 313-variant polygenic risk score (PRS) for BC in 6,339 women of European ancestry aged \geq 70 years. We evaluated incident BC diagnoses over a median follow-up of 4.7 years. A multivariable Cox regression model including conventional risk factors was applied to prospective data, and re-evaluated after adding the PRS. We also assessed the association of rare pathogenic variants (PVs) with BC in susceptibility genes (*BRCA1/BRCA2/PALB2/CHEK2/ATM*).

Results: The PRS, as a continuous variable, was an independent predictor of incident BC (hazard ratio [HR] per standard deviation (SD)=1.4, 95% confidence interval [CI] 1.3-1.6, N=110 cases) and hormone receptor (ER/PR)-positive disease (HR=1.5 [CI 1.2-1.9], N=79 cases). Women in the top quintile of the PRS distribution had higher risk of BC than women in the lowest quintile (HR=2.2 [CI 1.2-3.9]). The concordance index of the model without the PRS was 0.62 (95% CI 0.56-0.68) which improved after addition of the PRS to 0.65 (95% CI 0.59-0.71). Among 41 (0.6%) carriers of PVs in BC susceptibility genes, we observed no incident BC diagnoses.

Conclusion: The 313-variant PRS predicts BC risk in women aged 70 years and older.

1 Introduction

Breast cancer (BC) risk prediction models may be improved by including genomic risk scores. A
polygenic risk score (PRS) aggregates the effect of many common BC risk-associated variants into a
single measure.¹⁻⁴ Predictive performance of PRS for BC has mostly been assessed in women aged
40-69 years.^{1,3,5-9} The PRS performance, in terms of risk prediction in older women (aged ≥70 years)
is unclear, despite a high proportion of BC diagnoses occurring in this age group. It is also unclear the
predictive performance of a PRS for BC attenuates with age. Given the emerging clinical utility of
PRS for BC risk prediction and stratification, this requires further assessment.

A small proportion of women (<5%) carry rare pathogenic variants (PVs) in BC predisposition genes,
including *BRCA1*, *BRCA2*, *PALB2*, *CHEK2* and *ATM*.¹⁰ Rare PVs can be detected by predictive
clinical genetic testing and are of high clinical significance,^{11,12} predisposing women to BC earlier in
life^{5-7,10}. Rare PVs account for ~25% of familial BC risk.¹³. If rare PVs have age-dependent effects on
BC risk, this may have important clinical implications for the appropriateness of offering predictive
genetic testing to older people. Common BC risk-associated genetic variants used in a PRS, together,
are estimated to account for a further 18% of familial risk.¹⁴

Many clinical studies have measured the association between rare PVs and BC risk.^{5-7,10} More recently, risk conferred by a 313-variant PRS for BC has been measured in meta-analysis of ten studies, with no evidence of age-related attenuation in predictive performance of the PRS reported.¹ The same PRS has been assessed across several subsequent PRS validation studies.²⁻⁴ However, participant numbers aged \geq 70 years were limited. Here, we evaluate the predictive performance of i) a PRS for BC and ii) rare PVs in a prospective cohort of 6,339 women aged \geq 70 years.

22 <u>Methods</u>

23 Study sample

24 The study sample comprised female participants of the ASPirin in Reducing Events in the Elderly 25 (ASPREE) trial - a randomized, placebo-controlled, clinical trial investigating the effect of aspirin on disability-free survival.¹⁵⁻¹⁷. Study design^{18,19}, recruitment²⁰, and baseline characteristics²¹ have been 26 27 published previously. The trial recruited 19,114 individuals from Australia (N=16,703) and the United 28 States (N=2,411) aged 70 years or older (\geq 65 years for US ethnic minorities), who at enrolment were 29 free from diagnosed cardiovascular disease events, dementia, physical disability and life-threatening 30 cancer diagnoses. Biospecimens and consent for genetic analysis were obtained from 14,576 31 participants. The median follow-up time (randomization period) was 4.7 years (interquartile range 2.1 32 years). The study received local ethics committee approvals and is registered (NCT01038583).

33 Genome-wide genotyping and polygenic risk score

34 DNA samples were genotyped using the Axiom 2.0 Precision Medicine Diversity Research Array 35 (Thermo Fisher Scientific (TFS), CA, USA) following standard protocols. Only participants with 36 European genetic ancestry (>95% of female participants) were included, to mitigate population stratification bias. Genetic ancestry was defined using principal component analysis (PCA) based on 37 38 the 1,000 Genomes reference population, with participants outside of the Non-Finnish European ancestry cluster excluded (Figure S1).²². Imputation was performed using the TopMED Server 39 (European samples).²³ After variants with low imputation quality scores ($r^2 < 0.3$) were excluded, a 40 41 PRS was calculated based on the 313-variant score previously described (1), using genotypes for the remaining 271 variants.¹ Using Plink version 1.9, we calculated the PRS for each individual as the 42 43 weighted sum of the effect size for the number of risk alleles at each variant.¹

44 Targeted gene panel sequencing

Our custom gene panel²⁴ included BC predisposition genes that are incorporated into the Breast and
Ovarian Analysis of Disease Incidence and Carrier Estimation Algorithm (BOADICEA) - *BRCA1*, *BRCA2, PALB2, CHEK2 and ATM.*¹⁰ Following standard protocols, DNA was extracted and

sequenced using the S5TM XL system, to average 200X depth. Variants with 'pathogenic' or 'likely
pathogenic' ClinVar annotation²⁵ and/or high-confidence predicted loss-of-function in coding
regions²⁶ were curated following ACMG/AMP Standards and Guidelines for the Interpretation of
Sequence Variants,²⁷ including review by two or more laboratory scientists and a clinical geneticist.
Analysis was restricted to single nucleotide variants and small insertions/deletions.

53 Endpoints

The study's primary endpoint was invasive breast cancer (BC), which included incident invasive BC diagnosed during the ASPREE trial; this was adjudicated by an expert panel using histopathology, metastasis imaging or other clinical evidence.²⁸ Age at diagnosis of prevalent BC was self-reported as before or after 50 years.

58 Statistical analysis

59 Multivariable Cox proportional hazards regression was used to evaluate the association between PRS 60 with incident BC by estimating the hazard ratio (HR) per standard deviation (SD) of the PRS, after 61 adjusting for BC family history (first-degree blood relatives), treatment (aspirin/placebo), age at 62 enrolment, number of children, alcohol consumption, body mass index (BMI) at enrolment, and use of 63 estrogen or estrogen/progesterone hormone replacement therapy (HRT) at enrolment. Alcohol 64 consumption was categorised into three groupings: none (no current consumption); low (<3 drinking 65 days per week); and high (\geq 3 drinking days per week). Interaction between PRS and aspirin treatment 66 was tested independently alongside the complete set of covariates. BMI and number of children were 67 standardised to mean=0 (SD 1).

In a separate model, PRS was categorized into three quintile-based distribution groups -low (0-20%, Q1), moderate-risk (21-80%, Q2-4), and high-risk (81-100%, Q5). R package *pec* v2019.11.03 was used for BC-free survival prediction. Net reclassification improvement (NRI) was calculated using R package *nricens* v1.6, with 1% and 3% cutoff values for predicting increased and decreased risk categories. We calculated cumulative incidence estimates for each PRS group (multivariate adjusted), treating death as a competing risk.

Logistic regression was used to assess associations with prevalent breast cancer, including family 74 75 history and presence of rare PVs as covariates. We further stratified by diagnosis age (<50, 50+), and 76 included number of children in the model. We used variance inflation factor (VIF) to assess the independence of predictors, and measured the discriminative ability of the PRS using concordance 77 78 index and area under the receiver operating characteristic curve (AUC). Goodness-of-fit for the logistic regression was assessed using the Hosmer-Lemeshow (HL) test and the Tail-Based Max-test-79 statistic (TBM).²⁹ DeLong's test was used to compare between two correlated ROC curves.^{30,31} We 80 81 calculated PV ORs for prevalent BC for each gene and all five genes combined. Analyses were 82 performed using R v3.6.1.

83 <u>Results</u>

84 Baseline characteristics

- The mean age of the 6,339 female participants of European genetic ancestry was 75.1 years at time of
- enrolment, with 14% aged >80 years, and 31% current or former smokers (Table 1). The mean BMI
- was 28.0 kg/m^2 , 75% were current alcohol consumers and 13% had a family history of BC in a first-
- degree blood relative. At baseline, 533 (8.4%) participants were taking HRT (either estrogen alone or
- 89 with progesterone). Thirteen (0.2%) were taking progesterone-only preparations. Prevalent BC, was
- 90 reported by 475 (7.6%) participants, of which 60 (1%) were diagnosed before the age of 50 years.

91 PRS and rare pathogenic variants

The PRS showed a normal distribution in the study sample with mean -0.13 (SD 0.58) before standardization (Figure S2a), which was scaled to a mean of 0 (SD 1) for subsequent analyses. Fortytwo rare pathogenic or likely pathogenic variants (PVs) passed our variant curation protocol (Table S1) across 41 (0.6%) participants (one participant had two PVs detected, one each in *BRCA2* and *CHEK2*). We identified participants with PVs in the *BRCA1* (N=3), *BRCA2* (N=10), *PALB2* (N=6), *CHEK2* (N=7) and *ATM* (N=16) genes. Of these participants, 20% reported a family history of BC in a first-degree blood relative at enrolment.

99 Incident breast cancer risk

During median follow-up (4.7 years/participant), 110 women had incident BC, with no prevalent BC.

101 None of these women had rare PVs in the BRCA1, BRCA2, PALB2, CHEK2, and ATM genes. In the

102 multivariable Cox model, conventional BC risk factors, including family history of BC, number of

- 103 children, alcohol consumption and estrogen HRT were associated with risk of incident BC (Table 2).
- 104 The PRS, as a continuous variable in the same model, was an independent predictor of incident BC,
- 105 with a HR of 1.43 (95% confidence interval [CI] 1.18 to 1.73, p<0.001) per SD (Table 2, Figure 1),
- after adjustment for covariates. The VIF for each term in the multivariable model was less than 1.1,
- 107 indicating independence of the predictors. The concordance index of the model without the PRS was

108 0.62 (95% CI 0.56 to 0.68), which improved after addition of the PRS to 0.65 (95% CI 0.59 to 0.71).

109 We found no evidence of an interaction between aspirin treatment and the PRS.

110 We categorized the PRS into low- (Q1), moderate- (Q2-4) and high-risk (Q5) groups to consider PRS 111 effect on incident BC. When using Q1 as a reference, participants in the high-risk PRS group had a 112 significantly higher risk of developing incident BC compared to women in the low-risk PRS group 113 (HR=2.16 [95% CI 1.21 to 3.86], p<0.01) (Table 2, Figure 2). The competing risk model showed that 114 individuals in Q5 (the high-risk group) had higher cumulative incidence than those in Q1 (the low-risk 115 group) and Q2-4 (the moderate-risk group) (Figure 3). Participants in the moderate- and low-risk 116 groups did not have significantly different risks of incident BC. The calibration plot for the incident 117 risk model (Figure S2b), illustrates high concordance between the predicted and observed events. Net 118 reclassification analysis had point estimates of 0.15 [95% CI 0.03; 0.24] for combined change, with 119 NRI+ 0.09 [95% CI -0.02; 0.13] and NRI- of 0.05 [95% CI 0.02; 0.08]. Reclassification of cases and 120 controls is shown in Table S2.

Histopathology was available for 103 incident BC cases (Table S3). The PRS was found to be a
significant predictor of (ER+/PR+) disease (HR=1.53 per SD [95% CI 1.22 to 1.91)], p<0.001, n=79).

123 Prevalent breast cancer

Of the 41 participants with PVs, 11 (27.5%) reported prevalent BC at baseline, compared with 7.5% (475/6,339) in all female participants, giving an estimated OR of 4.69 (95% CI 2.21 to 9.27, p<0.001) for PVs grouped across all five genes (Table S4). The OR estimate for PVs in *BRCA1* and *BRCA2* (OR=6.09 [95% CI 1.77 to 19.08]) was higher than PVs in the other genes (*ATM/PALB/CHECK2*) (OR=3.33 [95% CI 1.19 to 7.92, p=0.011). Per-gene ORs are reported in Table S5, but are limited by small carrier numbers.

130 In sub-group analysis, when stratifying prevalent BC cases by diagnosis age (before or after 50 years),

the OR estimate for having a PV was higher for early-onset BC risk (diagnosed <50 years, OR=9.79

132 [CI 2.29 to 28.87], p=0.02) than for later-onset BC risk (diagnosed >50 years, OR=3.91 [CI 1.64 to

133 8.31], p=0.02) (Table S6).

134 The PRS as a continuous variable was associated with prevalent BC after controlling for covariates in 135 the model (OR=1.47 per SD [95% CI 1.34 to 1.61, p<0.001). The HL and TBM tests did not indicate 136 a lack of goodness-of-fit (p>0.05). The AUC for the model with the PRS (AUC=0.62 [0.59-0.65], 137 Figure S2c) was improved relative to the model including family history of BC only (AUC=0.53) 138 [0.52-0.55], Figure S2d) (p<0.01). When considering the PRS as a categorical variable, participants in 139 the high-risk group had a significantly higher BC risk compared with the low-risk group (OR=3.16 140 [95% CI 2.26 to 4.49], p<0.001). Participants in the moderate-risk group also had higher BC risk 141 versus the low-risk group (OR=2.12 [95% CI 1.56 to 2.94], p<0.001.). 142 Modification of BC risk by PRS in rare PV carriers

Eleven participants with PVs reported prevalent BC, and 29 with PVs reported no prevalent BC. We hypothesized that individuals with PVs and a history of BC (affected) may have a higher PRS, on average, than those with PVs but without BC (unaffected), as suggested by previous *BRCA1/BRCA2* studies.³² However, we observed no evidence of over-representation of affected or unaffected carriers between low-, moderate- or high-risk PRS groups (Chi-squared χ 2=1.97, df=2, P=0.37) (Figure 3, Table S7).

149 Discussion

150 The 313-variant PRS for BC predicted incident BC risk in the ASPREE cohort, for women aged \geq 70 151 years as both a continuous (per SD) or categorical (low, moderate and high-risk groups) variable, and 152 improved net reclassification when added to a model composed of traditional BC risk factors. This 153 suggests adding a PRS may increase predictive performance above traditional clinical risk factors for 154 models predicting BC risk in older women. Women in the highest quintile of the PRS distribution had 155 over 2-fold higher risk of incident BC than those in the lowest quintile. The PRS was also associated 156 with incident hormone receptor positive (ER+/PR+) BC specifically (HR=1.5 [95% CI 1.2-1.9]) and 157 with prevalent BC (HR=1.4 per SD [95% CI 1.3-1.6]). The emerging clinical utility of PRS for BC risk prediction and risk stratification, previously demonstrated in women aged 40-69 years ¹⁻⁴, 158 159 therefore likely extends to older women.

160 The PRS effect (HR=1.4 per SD) was similar to that reported in population-based studies of younger women,^{1,3,8,9} including a recent meta-analysis measuring the same PRS used in our study across ten 161 162 prospective studies of younger women where ORs ranged from 1.48 to 1.75 across participants of all 163 ages.¹ Subsequent validation studies of the same PRS in other cohorts found similar ORs to our study (~ 1.4) , with no evidence of age-related attenuation.²⁻⁴ The average age at enrolment in ASPREE (75 164 165 years) is over 15 years older than other population-based studies, including the UK Biobank, where 166 average enrolment age was 58 years³³, yet similar HRs were observed. However, HR=1.4 is a modest 167 effect. Most incident BC cases in our study had favourable prognoses (e.g. hormone receptor 168 positive), meaning the PRS impact for improving survival in older women in a clinical setting may be 169 limited, and must be balanced against overdiagnosis/overtreatment risks in this demographic.

170 No incident BC diagnoses were observed in 41 women with rare PVs, despite the high expected BC 171 risk conferred by these PVs (*e.g.* average cumulative risk to age 70 years of 50-70% for 172 *BRCA1/BRCA2* PVs).⁷ This challenges the clinical value of predictive genetic testing for BC risk by 173 sequencing of these genes alone in women aged \geq 70 years. However, this testing can trigger cascade 174 family testing, which has benefits including risk management, early detection, and/or prevention of

175 cancer in younger family members. We also observed no incident ovarian cancer diagnoses in the 41

176 PV carriers.

177 Retrospective data suggested a higher risk of BC in individuals with a PV compared with those 178 without (OR=4.7) when all PVs across all genes were combined into a single group. This reflects the 179 effect of PVs earlier in life. However, women with PVs diagnosed with cancer earlier in life are less 180 likely to have been ascertained by our study, because they either died from cancer before the 181 enrolment age or were too unwell to enrol in the ASPREE trial due to a current or recent cancer 182 diagnosis. This has likely resulted the healthy selection bias that can often occur in older survivor 183 cohorts.³⁴ Thus risk estimates observed in PV carriers ascertained in our study (and the ORs and CIs 184 reported for associations between the PRS/PVs and prevalent cases) must be interpreted with caution.

In response to recent studies reporting that an individual's PRS may modify the penetrance of rare PVs in the *BRCA1* and *BRCA2* genes,^{8,9,32,35} we sought evidence of risk-modification by the PRS in PV carriers affected and unaffected by prevalent BC in our study. We observed no evidence of either a higher PRS in females with a PV and affected by BC, or a protective effect of a lower PRS in unaffected PV carriers (Figure 3). However, we acknowledge our analysis was limited by a relatively small number of PV carriers. Further studies are needed to investigate this more rigorously.

Key strengths of our study include the well-characterised, older study population (median age after follow-up 78 years) followed prospectively, with all incident BC diagnoses adjudicated by an expert panel. Most previous studies of genetic risk scores for BC have examined younger cohorts, some selected for family history.

Limitations of our study include the unavailability of some phenotypic and clinical risk factors associated with BC, such as mammographic density, reproductive factors (*e.g.* age at menarche, menopause, first birth) and hormonal factors beyond HRT use (*e.g.* oral contraceptive use)³⁶. Our clinical risk factor model might be improved with these additional factors. Participants may have undergone risk-reducing bilateral prophylactic mastectomy or bilateral salpingo-oophorectomy prior to enrolment, to reduce BC/OC risk. The relatively small number of PVs detected (N=41) necessitated grouping of PVs across genes to calculate averaged risk, despite known differences in the magnitude

of risk conferred by PVs in different genes.^{11,12} The relatively short follow-up period of 4.7 years limited the total number of incident BC events. Prevalent BC (pre-enrolment) was self-reported (without specific diagnosis age) and not verified through supporting documentation, potentially causing over-estimation of prevalent BC events. Our study involved only participants of European genetic ancestry, meaning results may not be generalisable to other populations.

207 In conclusion, we demonstrate that the predictive value of a PRS extends to older women, with no

208 evidence of age-related attenuation in predictive performance after age 70 years. Our study has

209 clinical implications for the use and interpretation of polygenic risk prediction of BC across the

210 female lifespan.

211 <u>Notes</u>:

212 **Data availability:** The data underlying this article will be shared on reasonable request to the 213 corresponding author.

214 Funding: This work was supported by an ASPREE Flagship cluster grant (including the 215 Commonwealth Scientific and Industrial Research Organisation, Monash University, Menzies 216 Research Institute, Australian National University, University of Melbourne); and grants 217 (U01AG029824 and U19AG062682) from the National Institute on Aging and the National Cancer 218 Institute at the National Institutes of Health, by grants (334047 and 1127060) from the National 219 Health and Medical Research Council of Australia, and by Monash University and the Victorian 220 Cancer Agency. AEC receives a NHMRC Career Development Fellowship (1147843). HPS holds an 221 NHMRC MRFF Next Generation Clinical Researchers Program Practitioner Fellowship 222 (APP1137127). PL is supported by a National Heart Foundation Future Leader Fellowship.

223 Author disclosures: No conflicts are declared.

224 Author contributions: PL: Conceptualization, Data curation, Funding acquisition, Methodology, 225 Supervision, Writing – original draft. AB: Data curation, Formal Analysis, Investigation, Software. 226 MR: Conceptualization, Data curation, Formal Analysis, Investigation, Software. SGO: Resources, 227 Data curation, Investigation, Project administration. JT: Investigation, Writing – original draft. JN: 228 Data curation, Formal Analysis, Investigation. PRC: JN: Data curation, Formal Analysis, 229 Investigation. AJ: Conceptualization, Methodology. YC: Conceptualization, Methodology. ETW: 230 Conceptualization, Validation. AM: Methodology, Validation. TND: Conceptualization, Validation. 231 MCSS: Conceptualization, Validation. RLM: Conceptualization, Validation. LFL: Conceptualization, 232 Data curation. RS: Conceptualization, Resources, Data curation. Methodology. ES: 233 Conceptualization, Resources, LG: Data curation, Methodology, PG: Conceptualization, Resources, 234 Data curation, Methodology. BAT: Resources, Data curation, Formal Analysis, Methodology. FAM: 235 Conceptualization, Data curation. IW: Conceptualization, Resources, Data curation. CM: 236 Conceptualization, Resources, Data curation, Methodology, JRZ: Conceptualization, Resources, 237 Methodology. RW: Conceptualization, Resources, Methodology. ATC: Resources, Funding

- 238 acquisition, Project administration, Supervision. AM: Resources, Funding acquisition, Project
- administration. JJM: Conceptualization, Resources, Project administration, Supervision. All authors
- 240 were involved in writing (review and editing).
- 241 Acknowledgements: We thank the trial staff in Australia and the United States, the participants who
- volunteered for this trial, and the general practitioners and staff of the medical clinics who cared for
- the participants.

	Total	Low-risk PRS (Q1)	Moderate- risk PRS (Q2-4)	High-risk PRS (Q5)		
Participants	N = 6339	N = 1268	N = 3803	N = 1268		
Sex = Female (%)	6339 (100)	1268 (100)	3803 (100)	1268 (100)		
Mean Age at Enrolment, Years	75.1	75.1	75.1	75.2		
Age Group, Years (%)						
70-74	3825 (60.3)	778 (61.4)	2287 (60.1)	760 (59.9)		
75-79	1599 (25.5)	304 (24.0)	975 (25.6)	320 (25.2)		
80-84	706 (11.1)	136 (10.7)	429 (11.3)	141 (11.1)		
85+	211 (3.3)	50 (3.9)	112 (2.9)	47 (3.7)		
Current or former smoker (%)	1974 (31.1)	391 (30.8)	1195 (31.4)	388 (30.6)		
Diabetes (%)	497 (7.8)	97 (7.6)	289 (7.6)	111 (8.8)		
Randomized to Aspirin (%)	3170 (50.0)	630 (49.7)	1898 (49.9)	642 (50.6)		
Body-mass-index kg/m ² (mean) (SD))	28.03 (5.09)	28.02 (5.02)	28.05 (5.10)	28.00 (5.13)		
Current alcohol consumption (%)	4730 (74.6)	941 (74.2)	2850 (74.9)	939 (74.1)		
Hormone replacement therapy*	533 (8.4)	103 (8.1)	321 (8.4)	109 (8.6)		
Progesterone-only HRT	13 (0.2)	1 (0.1)	3 (0.1)	2 (0.2)		
Family history of breast cancer (%)□	850 (13.4)	135 (10.6)	500 (13.1)	215 (17.0)		
Prevalent Breast Cancer	•	•	•	•		
Cases	475	47	288	140		
Diagnosed < 49 Years	60	6	39	15		
Diagnosed 50+ Years	415	41	249	125		
Incident Breast Cancer¶	Incident Breast Cancer					
Cases	129	21	66	42		
Polygenic Risk Score (mean (SD))	0.1 (0.53)	-0.93 (0.26)	-0.13 (0.27)	0.69 (0.28)		

Table 1. Characteristics of the study population

PRS = Polygenic risk score, Q = Quintile, HRT = Hormone replacement therapy

* Estrogen alone or in combination with progesterone.

¶ Non-metastatic and metastatic events

□ Family history in first-degree blood relative (mother, sibling or child).

Table 2: Association of a polygenic risk score (PRS) with incident breast cancer (BC) risk in 6,339 older women. A multivariable Cox proportional hazards regression model was used to evaluate the association between PRS as a continuous or categorical variable with incident BC (N=110 cases), after adjusting for family history of BC, treatment (aspirin/placebo), age at enrolment, number of children, alcohol use, BMI and use of hormone replacement therapy.

	PRS as Continuous Variable			PRS as Categorical Variable				
	Hazard Ratio	95% CI	p-value	Hazard Ratio	95% CI	p-value		
Polygenic Score (per standard deviation)	1.43	(1.18; 1.73)	< 0.001					
Low PRS (Q1)					Reference			
Moderate PRS (Q2,3,4)				1.16	(0.68; 2.00)	0.58		
High PRS (Q5)				2.16	(1.21; 3.86)	0.009		
Pathogenic variants (N=41 carriers)	No incident events			No incident events				
Family History of Breast Cancer* (Y/N)	1.81	(1.15; 2.85)	0.01	1.83	(1.16; 2.88)	0.009		
Age at Enrolment	0.97	(0.92; 1.02)	0.21	0.97	(0.92; 1.02)	0.22		
Treatment (Aspirin)	1.16	(0.80; 1.69)	0.44	1.15	(0.79; 1.68)	0.45		
Number of Children	0.81	(0.66; 0.99)	0.04	0.81	(0.66; 0.99)	0.04		
Body-mass-index (kg/m ² (mean) SD)	1.14	(0.95; 1.37)	0.17	1.14	(0.95; 1.37)	0.15		
Alcohol (None)		Reference			Reference			
Alcohol (Low)	1.16	(0.68; 1.97)	0.59	1.16	(0.68; 1.98)	0.58		
Alcohol (High)	1.70	(1.01; 2.85)	0.04	1.70	(1.02; 2.86)	0.04		
HRT (Y/N)	1.54	(0.88; 2.71)	0.13	1.51	(0.86; 2.65)	0.15		

BMI=Body mass index, HRT=Hormone replacement therapy, PRS=Polygenic risk score, CI=Confidence Interval, SD=Standard deviation

*Family history in first-degree blood relative (mother, sibling or child).

 $\hfill\square$ Estrogen alone or in combination with progesterone.

References

244	1.	Mavaddat N, Michailidou K, Dennis J, et al. Polygenic Risk Scores for Prediction of Breast
245		Cancer and Breast Cancer Subtypes. Am J Hum Genet. 2019;104(1):21-34.
246	2.	Lakeman IMM, Rodriguez-Girondo M, Lee A, et al. Validation of the BOADICEA model
247		and a 313-variant polygenic risk score for breast cancer risk prediction in a Dutch prospective
248		cohort. Genet Med. 2020.
249	3.	Kapoor PM, Mavaddat N, Choudhury PP, et al. Combined associations of a polygenic risk
250		score and classical risk factors with breast cancer risk. J Natl Cancer Inst. 2020.
251	4.	Mars N, Koskela JT, Ripatti P, et al. Polygenic and clinical risk scores and their impact on
252		age at onset and prediction of cardiometabolic diseases and common cancers. Nat Med.
253		2020;26(4):549-557.
254	5.	Prevalence and penetrance of BRCA1 and BRCA2 mutations in a population-based series of
255		breast cancer cases. Anglian Breast Cancer Study Group. Br J Cancer. 2000;83(10):1301-
256		1308.
257	6.	Kuchenbaecker KB, Hopper JL, Barnes DR, et al. Risks of Breast, Ovarian, and Contralateral
258		Breast Cancer for BRCA1 and BRCA2 Mutation Carriers. JAMA. 2017;317(23):2402-2416.
259	7.	Mavaddat N, Peock S, Frost D, et al. Cancer risks for BRCA1 and BRCA2 mutation carriers:
260		results from prospective analysis of EMBRACE. J Natl Cancer Inst. 2013;105(11):812-822.
261	8.	Barnes DR, Rookus MA, McGuffog L, et al. Polygenic risk scores and breast and epithelial
262		ovarian cancer risks for carriers of BRCA1 and BRCA2 pathogenic variants. Genet Med.
263		2020;22(10):1653-1666.
264	9.	Kuchenbaecker KB, McGuffog L, Barrowdale D, et al. Evaluation of Polygenic Risk Scores
265		for Breast and Ovarian Cancer Risk Prediction in BRCA1 and BRCA2 Mutation Carriers. J
266		Natl Cancer Inst. 2017;109(7).
267	10.	Lee AJ, Cunningham AP, Tischkowitz M, et al. Incorporating truncating variants in PALB2,
268		CHEK2, and ATM into the BOADICEA breast cancer risk model. Genet Med.
269		2016;18(12):1190-1198.

18

- 270 11. Breast Cancer Association C, Dorling L, Carvalho S, et al. Breast Cancer Risk Genes -
- Association Analysis in More than 113,000 Women. N Engl J Med. 2021;384(5):428-439.
- Hu C, Hart SN, Gnanaolivu R, et al. A Population-Based Study of Genes Previously
 Implicated in Breast Cancer. *N Engl J Med.* 2021;384(5):440-451.
- Bahcall O. Common variation and heritability estimates for breast, ovarian and prostate
 cancers. *Nature Genetics*. 2013.
- Michailidou K, Lindstrom S, Dennis J, et al. Association analysis identifies 65 new breast
 cancer risk loci. *Nature*. 2017;551(7678):92-94.
- McNeil JJ, Wolfe R, Woods RL, et al. Effect of Aspirin on Cardiovascular Events and
 Bleeding in the Healthy Elderly. *N Engl J Med.* 2018;379(16):1509-1518.
- McNeil JJ, Woods RL, Nelson MR, et al. Effect of Aspirin on Disability-free Survival in the
 Healthy Elderly. *N Engl J Med.* 2018;379(16):1499-1508.
- McNeil JJ, Nelson MR, Woods RL, et al. Effect of Aspirin on All-Cause Mortality in the
 Healthy Elderly. *N Engl J Med.* 2018;379(16):1519-1528.
- 284 18. Group AI. Study design of ASPirin in Reducing Events in the Elderly (ASPREE): a
 285 randomized, controlled trial. *Contemporary clinical trials*. 2013;36(2):555-564.
- 286 19. Nelson MR, Reid CM, Ames DA, et al. Feasibility of conducting a primary prevention trial of
- low-dose aspirin for major adverse cardiovascular events in older people in Australia: results
 from the ASPirin in Reducing Events in the Elderly (ASPREE) pilot study. *The Medical journal of Australia*. 2008;189(2):105-109.
- 20. Lockery JE, Collyer TA, Abhayaratna WP, et al. Recruiting general practice patients for large
 clinical trials: lessons from the Aspirin in Reducing Events in the Elderly (ASPREE) study. *Med J Aust.* 2019;210(4):168-173.
- 293 21. McNeil JJ, Woods RL, Nelson MR, et al. Baseline Characteristics of Participants in the
 294 ASPREE (ASPirin in Reducing Events in the Elderly) Study. *J Gerontol A Biol Sci Med Sci*.
 295 2017;72(11):1586-1593.
- 296 22. Genomes Project C, Auton A, Brooks LD, et al. A global reference for human genetic
 297 variation. *Nature*. 2015;526(7571):68-74.

- 298 23. Taliun D, Harris DN, Kessler MD, et al. Sequencing of 53,831 diverse genomes from the
- 299 NHLBI TOPMed Program. *Nature*. 2021;590(7845):290-299.
- Lacaze P, Sebra R, Riaz M, et al. Medically actionable pathogenic variants in a population of
 13,131 healthy elderly individuals. *Genet Med.* 2020.
- Landrum MJ, Lee JM, Riley GR, et al. ClinVar: public archive of relationships among
 sequence variation and human phenotype. *Nucleic acids research*. 2014;42(Database
 issue):D980-985.
- Lek M, Karczewski KJ, Minikel EV, et al. Analysis of protein-coding genetic variation in
 60,706 humans. *Nature*. 2016;536(7616):285-291.
- Richards S, Aziz N, Bale S, et al. Standards and guidelines for the interpretation of sequence
 variants: a joint consensus recommendation of the American College of Medical Genetics and
- 309 Genomics and the Association for Molecular Pathology. *Genet Med.* 2015;17(5):405-424.
- Orchard SG, Lockery JE, Gibbs P, et al. Cancer history and risk factors in healthy older
 people enrolling in the ASPREE clinical trial. *Contemp Clin Trials*. 2020;96:106095.
- Song M, Kraft P, Joshi AD, Barrdahl M, Chatterjee N. Testing calibration of risk models at
 extremes of disease risk. *Biostatistics*. 2015;16(1):143-154.
- 30. DeLong ER, DeLong DM, Clarke-Pearson DL. Comparing the areas under two or more
 correlated receiver operating characteristic curves: a nonparametric approach. *Biometrics*.
 1988;44(3):837-845.
- 317 31. Sun X, Xu W. Fast Implementation of DeLong's Algorithm for Comparing the Areas Under
 318 Correlated Receiver Operating Characteristic Curves. *IEEE Signal Processing Letters*.
 319 2014;21(11):1389-1393.
- 320 32. Fahed AC, Wang M, Homburger JR, et al. Polygenic background modifies penetrance of
 321 monogenic variants for tier 1 genomic conditions. *Nat Commun.* 2020;11(1):3635.
- 322 33. Sudlow C, Gallacher J, Allen N, et al. UK biobank: an open access resource for identifying
 the causes of a wide range of complex diseases of middle and old age. *PLoS Med.*2015;12(3):e1001779.

20

- 325 34. Stovitz SD, Banack HR, Kaufman JS. 'Depletion of the susceptibles' taught through a story, a
- table and basic arithmetic. *BMJ Evid Based Med.* 2018;23(5):199.
- 327 35. Khera AV, Chaffin M, Aragam KG, et al. Genome-wide polygenic scores for common
 328 diseases identify individuals with risk equivalent to monogenic mutations. *Nat Genet*.
 329 2018;50(9):1219-1224.
- 330 36. 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(8):1708-1718.

Figure legends:

Figure 1. Association of a polygenic risk score (PRS) with incident and prevalent breast cancer (BC) risk in 6,339 older women. We evaluated incident BC diagnoses over a median follow-up of 4.7 years and prevalent BC diagnosed pre-enrolment (self-reported). A multivariable Cox regression model including conventional risk factors examined association between incident BC risk and the PRS as a categorical variable by quintiles (Q) of the distribution (low-[Q1] medium-[Q2-4], high-[Q5] risk groups). Logistic regression examined associations with prevalent BC.



Figure 2: Competing risk survival curves for incident breast cancer according to PRS groups.

The PRS distribution was categorized by quintiles (Q) of the distribution into three groups; low-risk (Q1, green), medium-risk (Q2-4, brown) and high-risk (Q5, red). Competing risk estimates of the cumulative incidence were calculated for each group (multivariable adjusted).



Figure 3: PRS distribution of female pathogenic variant (PV) carriers who were affected (red) or unaffected (blue) by prevalent BC. Density represents the proportion of individuals in each PRS group (low/medium/high). The PRS is distributed normally according to scaled PRS score (mean 0, standard deviation 1). PV carriers are highlighted across the PRS distribution.

