Overlap of high-risk individuals across family history, genetic & non-genetic breast cancer risk models: Analysis of 180,398 women from European & Asian ancestries.

RUNNING TITLE

Overlap of high-risk predictions in breast cancer models

Peh Joo Ho (PhD)^{1,2}, Christine Kim Yan Loo¹, Meng Huang Goh¹, Mustapha Abubakar³, Thomas U. Ahearn (PhD)³, Irene L. Andrulis (PhD)^{4,5}, Natalia N. Antonenkova (MD)⁶, Kristan J. Aronson (PhD)⁷, Annelie Augustinsson (PhD)⁸, Sabine Behrens⁹, Clara Bodelon¹⁰, Natalia V. Bogdanova (PhD)^{6,11,12}, Manjeet K. Bolla (MSc)¹³, Kristen Brantley¹⁴, Hermann Brenner (MD)^{15,16}, Helen Byers¹⁷, Nicola J. Camp (PhD)¹⁸, Jose E. Castelao (MD)¹⁹, Melissa H. Cessna (MD)²⁰, Jenny Chang-Claude (PhD)^{9,21}, Stephen J. Chanock (MD)³, Georgia Chenevix-Trench (PhD)²², Ji-Yeob Choi (PhD)²³⁻²⁵, Sarah V. Colonna (MD)¹⁸, Kamila Czene (PhD)²⁶, Mary B. Daly (MD)²⁷, Francoise Derouane²⁸, Thilo Dörk (PhD)¹², A. Heather Eliassen (ScD)^{14,29,30}, Christoph Engel (MD)^{31,32}, Mikael Eriksson (PhD)²⁶, D. Gareth Evans (MD) ^{17,33}, Olivia Fletcher (PhD) ³⁴, Lin Fritschi (PhD) ³⁵, Manuela Gago-Dominguez (MD) ³⁶, Jeanine M. Genkinger^{37,38}, Willemina R.R. Geurts-Giele³⁹, Gord Glendon (MSc)⁴, Per Hall (PhD)^{26,40}, Ute Hamann (PhD)⁴¹, Cecilia Y.S. Ho⁴², Weang-Kee Ho^{43,44}, Maartje J. Hooning (PhD) ^{39,45}, Reiner Hoppe (PhD)^{46,47}, Anthony Howell (MBBS)⁴⁸, Keith Humphreys (PhD)²⁶, ABCTB Investigators⁴⁹, kConFab Investigators^{50,51}, SGBCC Investigators^{1,2,52-62}, MyBrCa Investigators^{43,44,63}, Hidemi Ito^{64,65}, Motoki Iwasaki⁶⁶, Anna Jakubowska (PhD) ^{67,68}, Helena Jernström (PhD) ⁸, Esther M. John (PhD) ^{69,70}, Nichola Johnson (DPhil) ³⁴, Daehee Kang (MD) ^{24,71}, Sung-Won Kim (MD) ⁷², Cari M. Kitahara (PhD)⁷³, Yon-Dschun Ko (MD)⁷⁴, Peter Kraft (PhD)³, Ava Kwong⁷⁵⁻⁷⁷, Diether Lambrechts (PhD)^{78,79}, Susanna Larsson (PhD)^{80,81}, Shuai Li (PhD)^{13,82,83}, Annika Lindblom (MD)^{84,85}, Martha Linet (MD)⁷³, Jolanta Lissowska (PhD)⁸⁶, Artitaya Lophatananon (PhD)⁸⁷, Robert J. MacInnis (PhD) ^{82,88}, Arto Mannermaa (PhD) ⁸⁹⁻⁹¹, Siranoush Manoukian (MD) ⁹², Sara Margolin (MD) ^{40,93}, Keitaro Matsuo (MD)^{65,94}, Kyriaki Michailidou (PhD)^{13,95}, Roger L. Milne (PhD)^{82,83,88}, Nur Aishah Mohd Taib (MD)⁹⁶, Kenneth Muir (PhD)⁸⁷, Rachel A. Murphy (PhD)^{97,98}, William G. Newman^{17,33}, Katie M. O'Brien (PhD)⁹⁹, Nadia Obi (PhD)^{100,101}, Olufunmilayo I. Olopade¹⁰², Mihalis I. Panayiotidis (PhD)¹⁰³, Sue K. Park (MD)^{24,71,104}, Tjoung-Won Park-Simon (MD)¹², Alpa V. Patel (PhD)¹⁰, Paolo Peterlongo (PhD)¹⁰⁵, Dijana Plaseska-Karanfilska (PhD)¹⁰⁶, Katri Pylkäs (PhD)^{107,108}, Muhammad U. Rashid (MD)^{41,109}, Gad Rennert (MD)¹¹⁰, Juan Rodriguez (PhD)²⁶, Emmanouil Saloustros (MD)¹¹¹, Dale P. Sandler (PhD)⁹⁹, Elinor J. Sawyer (PhD)¹¹², Christopher G. Scott (MS)¹¹³, Shamim Shahi (PhD)⁹⁷, Xiao-Ou Shu (MD)¹¹⁴, Katerina Shulman (MD)¹¹⁵, Jacques Simard (PhD)¹¹⁶, Melissa C. Southey (PhD) ^{83,88,117}, Jennifer Stone (PhD) ^{82,118}, Jack A. Taylor (MD) ^{99,119}, Soo-Hwang Teo (PhD) ^{44,96}, Lauren R. Teras (PhD)¹⁰, Mary Beth Terry (PhD)³⁷, Diana Torres (PhD)^{41,120}, Celine M. Vachon (PhD) ¹²¹, Maxime Van Houdt²⁸, Jelle Verhoeven²⁸, Clarice R. Weinberg (PhD) ¹²², Alicja Wolk (DrMedSci)⁸⁰, Taiki Yamaji⁶⁶, Cheng Har Yip (MD)^{96,123}, Wei Zheng (MD)¹¹⁴, Mikael Hartman (MD) ^{2,52,53}. Jingmei Li (PhD) ^{1,53,124}

- 1 Human Genetics Division, *Genome Institute of Singapore, Agency for Science, Technology* and Research (A*STAR), Singapore, Singapore City, 138672.
- 2 Saw Swee Hock School of Public Health, *National University of Singapore*, Singapore, Singapore City, 117549.
- 3 Division of Cancer Epidemiology and Genetics, *National Cancer Institute, National Institutes* of Health, Department of Health and Human Services, USA, Bethesda, MD, 20850.
- 4 Fred A. Litwin Center for Cancer Genetics, *Lunenfeld-Tanenbaum Research Institute of Mount Sinai Hospital*, Canada, Toronto, Ontario, M5G 1X5.
- 5 Department of Molecular Genetics, *University of Toronto*, Canada, Toronto, Ontario, M5S 1A8.
- 6 *N.N. Alexandrov Research Institute of Oncology and Medical Radiology*, Belarus, Minsk, 223040.

- 7 Department of Public Health Sciences, and Sinclair Cancer Research Institute, *Queen's University*, Canada, Kingston, ON, K7L 3N6.
- 8 Oncology, Department of Clinical Sciences in Lund, *Lund University*, Sweden, Lund, 221 85.
- 9 Division of Cancer Epidemiology, *German Cancer Research Center (DKFZ)*, Germany, Heidelberg, 69120.
- 10 Department of Population Science, *American Cancer Society*, USA, Atlanta, GA, 30303.
- 11 Department of Radiation Oncology, *Hannover Medical School*, Germany, Hannover, 30625.
- 12 Gynaecology Research Unit, Hannover Medical School, Germany, Hannover, 30625.
- 13 Centre for Cancer Genetic Epidemiology, Department of Public Health and Primary Care, University of Cambridge, UK, Cambridge, CB1 8RN.
- 14 Department of Epidemiology, *Harvard T.H. Chan School of Public Health*, USA, Boston, MA, 02115.
- 15 Division of Clinical Epidemiology and Aging Research, *German Cancer Research Center* (*DKFZ*), Germany, Heidelberg, 69120.
- 16 German Cancer Consortium (DKTK), *German Cancer Research Center (DKFZ)*, Germany, Heidelberg, 69120.
- 17 Division of Evolution and Genomic Sciences, School of Biological Sciences, Faculty of Biology, Medicine and Health, *University of Manchester, Manchester Academic Health Science Centre*, UK, Manchester, M13 9WL.
- 18 Department of Internal Medicine and Huntsman Cancer Institute, *University of Utah*, USA, Salt Lake City, UT, 84112.
- 19 Oncology and Genetics Unit, Instituto de Investigación Sanitaria de Santiago de Compostela (IDIS) Foundation, Complejo Hospitalario Universitario de Santiago, SERGAS, Spain, Vigo, 36312.
- 20 Intermountain Healthcare, USA, Salt Lake City, UT, 84143.
- 21 Cancer Epidemiology Group, University Cancer Center Hamburg (UCCH), University Medical Center Hamburg-Eppendorf, Germany, Hamburg, 20246.
- 22 Cancer Research Program, *QIMR Berghofer Medical Research Institute* Vol. Locked Bag 2000, Herston, QLD 4029, Australia, Brisbane, Queensland, 4006.
- 23 Department of Biomedical Sciences, *Seoul National University Graduate School*, Korea, Seoul, 03080.
- 24 Cancer Research Institute, Seoul National University, Korea, Seoul, 03080.
- 25 Institute of Health Policy and Management, *Seoul National University Medical Research Center*, Korea, Seoul, 03080.
- 26 Department of Medical Epidemiology and Biostatistics, *Karolinska Institutet*, Sweden, Stockholm, 171 65.
- 27 Department of Clinical Genetics, *Fox Chase Cancer Center*, USA, Philadelphia, PA, 19111.
- 28 Leuven Multidisciplinary Breast Center, Department of Oncology, *Leuven Cancer Institute, University Hospitals Leuven*, Belgium, Leuven, 3000.
- 29 Channing Division of Network Medicine, Department of Medicine, *Brigham and Women's* Hospital and Harvard Medical School, USA, Boston, MA, 02115.
- 30 Department of Nutrition, *Harvard T.H. Chan School of Public Health*, USA, Boston, MA, 02115.
- 31 Institute for Medical Informatics, Statistics and Epidemiology, *University of Leipzig*, Germany, Leipzig, 04107.
- 32 LIFE Leipzig Research Centre for Civilization Diseases, *University of Leipzig*, Germany, Leipzig, 04103.
- 33 North West Genomics Laboratory Hub, Manchester Centre for Genomic Medicine, St Mary's Hospital, Manchester University NHS Foundation Trust, Manchester Academic Health Science Centre, UK, Manchester, M13 9WL.
- 34 The Breast Cancer Now Toby Robins Research Centre, *The Institute of Cancer Research*, UK, London, SW7 3RP.
- 35 School of Population Health, *Curtin University*, Australia, Perth, Western Australia, 6102.
- 36 Cancer Genetics and Epidemiology Group, Genomic Medicine Group, Fundación Instituto de Investigación Sanitaria de Santiago de Compostela (FIDIS), Complejo Hospitalario Universitario de Santiago, SERGAS, Spain, Santiago de Compostela, 15706.
- 37 Department of Epidemiology, Mailman School of Public Health, *Columbia University*, USA, New York, NY, 10032.
- 38 Herbert Irving Comprehensive Cancer Center, USA, New York, NY.

- 39 Department of Clinical Genetics, *Erasmus University Medical Center* Vol. P.O. Box 2040, 3000 CA, the Netherlands, Rotterdam, 3015 CN.
- 40 Department of Oncology, Södersjukhuset, Sweden, Stockholm, 118 83.
- 41 Molecular Genetics of Breast Cancer, *German Cancer Research Center (DKFZ)*, Germany, Heidelberg, 69120.
- 42 Department of Molecular Pathology, Hong Kong Sanatorium and Hospital, Hong Kong.
- 43 School of Mathematical Sciences, Faculty of Science and Engineering, *University of Nottingham Malaysia*, Malaysia, Semenyih, Selangor, 43500.
- 44 Cancer Research Malaysia, Malaysia, Subang Jaya, Selangor, 47500.
- 45 Department of Medical Oncology, *Erasmus MC Cancer Institute* Vol. P.O. Box 2040, 3000 CA, the Netherlands, Rotterdam, 3015 GD.
- 46 Dr. Margarete Fischer-Bosch-Institute of Clinical Pharmacology, Germany, Stuttgart, 70376.
- 47 University of Tübingen, Germany, Tübingen, 72074.
- 48 Division of Cancer Sciences, University of Manchester, UK, Manchester, M13 9PL.
- 49 Australian Breast Cancer Tissue Bank, Westmead Institute for Medical Research, *University of Sydney*, Australia, Sydney, New South Wales, 2145.
- 50 Research Department, Peter MacCallum Cancer Center, Australia, Melbourne, Victoria, 3000.
- 51 Sir Peter MacCallum Department of Oncology, *The University of Melbourne*, Australia, Parkville, Victoria, 3010.
- 52 Department of Surgery, *National University Health System*, Singapore, Singapore City, 119228.
- 53 Department of Surgery, Yong Loo Lin School of Medicine, *National University of Singapore*, Singapore, Singapore City, 119077.
- 54 Cancer Genetics Service, National Cancer Centre, Singapore, Singapore City, 169610.
- 55 Breast Department, *KK Women's and Children's Hospital*, Singapore, Singapore City, 229899.
- 56 SingHealth Duke-NUS Breast Centre, Singapore, Singapore City, 168753.
- 57 Department of General Surgery, *Tan Tock Seng Hospital*, Singapore, Singapore City, 308433.
- 58 Division of Surgery and Surgical Oncology, *National Cancer Centre*, Singapore, Singapore City, 169610.
- 59 Department of General Surgery, *Singapore General Hospital*, Singapore, Singapore City, 169608.
- 60 Division of Breast Surgery, Department of General Surgery, *Changi General Hospital*, Singapore, Singapore City, 529889.
- 61 Division of Radiation Oncology, National Cancer Centre, Singapore, Singapore City, 169610.
- 62 Division of Medical Oncology, National Cancer Centre, Singapore, Singapore City, 169610.
- 63 Breast Cancer Research Unit, University Malaya Cancer Research Institute, Faculty of Medicine, *University of Malaya*, Malaysia, Kuala Lumpur, 50603.
- 64 Division of Cancer Information and Control, *Aichi Cancer Center Research Institute*, Japan, Nagoya, 464-8681.
- 65 Division of Cancer Epidemiology, *Nagoya University Graduate School of Medicine*, Japan, Nagoya, 466-8550.
- 66 Division of Epidemiology, National Cancer Center Institute for Cancer Control, Japan, Tokyo, 104-0045.
- 67 Independent Laboratory of Molecular Biology and Genetic Diagnostics, *Pomeranian Medical University*, Poland, Szczecin, 171-252.
- 68 International Hereditary Cancer Center, Department of Genetics and Pathology, *Pomeranian Medical University in Szczecin*, Poland, Szczecin, 70-115.
- 69 Department of Epidemiology and Population Health, *Stanford University School of Medicine*, USA, Stanford, CA, 94305.
- 70 Department of Medicine, Division of Oncology, *Stanford Cancer Institute, Stanford University School of Medicine*, USA, Stanford, CA, 94304.
- 71 Department of Preventive Medicine, *Seoul National University College of Medicine*, Korea, Seoul, 03080.
- 72 Department of Surgery, *Daerim Saint Mary's Hospital*, Korea, Seoul, 07442.
- 73 Radiation Epidemiology Branch, Division of Cancer Epidemiology and Genetics, *National Cancer Institute*, USA, Bethesda, MD, 20892.
- 74 Department of Internal Medicine, Johanniter GmbH Bonn, *Johanniter Krankenhaus*, Germany, Bonn, 53177.

- 75 Hong Kong Hereditary Breast Cancer Family Registry, Hong Kong.
- 76 Department of Surgery, *The University of Hong Kong*, Hong Kong.
- 77 Department of Surgery and Cancer Genetics Center, *Hong Kong Sanatorium and Hospital*, Hong Kong.
- 78 Laboratory for Translational Genetics, Department of Human Genetics, *KU Leuven*, Belgium, Leuven, 3000.
- 79 VIB Center for Cancer Biology, VIB, Belgium, Leuven, 3001.
- 80 Institute of Environmental Medicine, Karolinska Institutet, Sweden, Stockholm, 171 77.
- 81 Department of Surgical Sciences, *Uppsala University*, Sweden, Uppsala, 751 05.
- 82 Centre for Epidemiology and Biostatistics, Melbourne School of Population and Global Health, *The University of Melbourne*, Australia, Melbourne, Victoria, 3010.
- 83 Precision Medicine, School of Clinical Sciences at Monash Health, *Monash University*, Australia, Clayton, Victoria, 3168.
- 84 Department of Molecular Medicine and Surgery, *Karolinska Institutet*, Sweden, Stockholm, 171 76.
- 85 Department of Clinical Genetics and Genomics, *Karolinska University Hospital*, Sweden, Stockholm, 171 76.
- 86 Department of Cancer Epidemiology and Prevention, *M. Sklodowska-Curie National Research Oncology Institute*, Poland, Warsaw, 02-034.
- 87 Division of Population Health, Health Services Research and Primary Care, School of Health Sciences, Faculty of Biology, Medicine and Health, *The University of Manchester*, UK, Manchester, M13 9PL.
- 88 Cancer Epidemiology Division, Cancer Council Victoria, Australia, Melbourne, Victoria, 3004.
- 89 Translational Cancer Research Area, *University of Eastern Finland*, Finland, Kuopio, 70210.
- 90 Institute of Clinical Medicine, Pathology and Forensic Medicine, *University of Eastern Finland*, Finland, Kuopio, 70210.
- 91 Biobank of Eastern Finland, *Kuopio University Hospital*, Finland, Kuopio.
- 92 Unit of Medical Genetics, Department of Medical Oncology and Hematology, *Fondazione IRCCS Istituto Nazionale dei Tumori di Milano*, Italy, Milan, 20133.
- 93 Department of Clinical Science and Education, Södersjukhuset, *Karolinska Institutet*, Sweden, Stockholm, 118 83.
- 94 Division of Cancer Epidemiology and Prevention, *Aichi Cancer Center Research Institute*, Japan, Nagoya, 464-8681.
- 95 Biostatistics Unit, *The Cyprus Institute of Neurology and Genetics* Vol. P.O.Box 23462, 1683, Nicosia, Cyprus, Cyprus, Nicosia, 2371.
- 96 Department of Surgery, Faculty of Medicine, University of Malaya, UM Cancer Research Institute, Malaysia, Kuala Lumpur, 50603.
- 97 School of Population and Public Health, *University of British Columbia*, Canada, Vancouver, BC, V6T 1Z4.
- 98 Cancer Control Research, BC Cancer Agency, Canada, Vancouver, BC, V5Z 1L3.
- 99 Epidemiology Branch, *National Institute of Environmental Health Sciences, NIH*, USA, Research Triangle Park, NC, 27709.
- 100 Institute for Occupational and Maritime Medicine, *University Medical Center Hamburg-Eppendorf*, Germany, Hamburg, 20246.
- 101 Institute for Medical Biometry and Epidemiology, *University Medical Center Hamburg-Eppendorf*, Germany, Hamburg, 20246.
- 102 Center for Clinical Cancer Genetics, *The University of Chicago*, USA, Chicago, IL, 60637.
- 103 Department of Cancer Genetics, Therapeutics and Ultrastructural Pathology, *The Cyprus Institute of Neurology & Genetics* Vol. P.O.Box 23462, 1683, Nicosia, Cyprus, Cyprus, Nicosia, 2371.
- 104 Integrated Major in Innovative Medical Science, Seoul National University College of Medicine, Korea, Seoul, 03080.
- 105 Genome Diagnostics Program, *IFOM ETS the AIRC Institute of Molecular Oncology*, Italy, Milan, 20139.
- 106 Research Centre for Genetic Engineering and Biotechnology 'Georgi D. Efremov', *MASA*, Republic of North Macedonia, Skopje, 1000.
- 107 Laboratory of Cancer Genetics and Tumor Biology, Translational Medicine Research Unit, Biocenter Oulu, *University of Oulu*, Finland, Oulu, 90220.
- 108 Laboratory of Cancer Genetics and Tumor Biology, *Northern Finland Laboratory Centre Oulu*, Finland, Oulu, 90220.

- 109 Department of Basic Sciences, *Shaukat Khanum Memorial Cancer Hospital and Research Centre (SKMCH & RC)*, Pakistan, Lahore, 54000.
- 110 Faculty of Medicine, Technion Israel Institute of Technology and Association for Promotion of Research in Precision Medicine, Israel, Haifa, 35254.
- 111 Division of Oncology, Faculty of Medicine, School of Health Sciences, *University of Thessaly*, Greece, Larissa, 411 10.
- 112 School of Cancer & Pharmaceutical Sciences, Comprehensive Cancer Centre, Guy's Campus, *King's College London*, UK, London.
- 113 Department of Quantitative Health Sciences, Division of Clinical Trials and Biostatistics, *Mayo Clinic*, USA, Rochester, MN, 55905.
- 114 Division of Epidemiology, Department of Medicine, Vanderbilt Epidemiology Center, Vanderbilt-Ingram Cancer Center, *Vanderbilt University School of Medicine*, USA, Nashville, TN, 37232.
- 115 Clalit Regional Oncology Unit, Haifa and Western Galilee District, Israel, Haifa.
- 116 Genomics Center, Centre Hospitalier Universitaire de Québec Université Laval Research Center, Canada, Québec City, Québec, G1V 4G2.
- 117 Department of Clinical Pathology, *The University of Melbourne*, Australia, Melbourne, Victoria, 3010.
- 118 Genetic Epidemiology Group, School of Population and Global Health, *University of Western Australia*, Australia, Perth, Western Australia, 6009.
- 119 Epigenetic and Stem Cell Biology Laboratory, *National Institute of Environmental Health Sciences, NIH*, USA, Research Triangle Park, NC, 27709.
- 120 Institute of Human Genetics, *Pontificia Universidad Javeriana*, Colombia, Bogota, 110231.
- 121 Department of Quantitative Health Sciences, Division of Epidemiology, *Mayo Clinic*, USA, Rochester, MN, 55905.
- 122 Biostatistics and Computational Biology Branch, *National Institute of Environmental Health Sciences, NIH*, USA, Research Triangle Park, NC, 27709.
- 123 Subang Jaya Medical Centre, Malaysia, Subang Jaya, Selangor, 47500.
- 124 National Cancer Centre, Singapore, Singapore City, 168583.

* Correspondence to:

Dr Jingmei Li, Genome Institute of Singapore (GIS), Agency for Science, Technology and Research

(A*STAR), 60 Biopolis Street, Singapore 138672, Singapore. Tel: (65) 6808 8312; Email: lijm1@gis.astar.edu.sg

-----<u>9</u>

Word count: 3569

Total number of tables: 4

Total number of figures: 3

ABSTRACT

Background

Breast cancer is multifactorial. Focusing on limited risk factors may miss high-risk individuals.

Methods

We assessed the performance and overlap of various risk factors in identifying high-risk individuals for invasive breast cancer (BrCa) and ductal carcinoma in situ (DCIS) in 161,849 European-ancestry and 18,549 Asian-ancestry women. Discriminatory ability was evaluated using the area under the receiver operating characteristic curve (AUC). High-risk criteria included: 5-year absolute risk $\geq 1.66\%$ by the Gail model [GAIL_{binary}]; first-degree family history of breast cancer [FH_{binary}]; 5-year absolute risk $\geq 1.66\%$ by a 313-variants polygenic risk score [PRS_{binary}]; and carriers of pathogenic variants in breast cancer predisposition genes [PTV_{binary}].

Findings

The 5-year absolute risk by PRS outperformed the Gail model in predicting BrCa (Europeans_{vs controls}: AUC_{PRS}=0.635 [0.632-0.638] vs AUC_{Gail}=0.492 [0.489-0.495]; Asians_{vs controls}: AUC_{PRS}=0.564 [0.556-0.573] vs AUC_{Gail}=0.506 [0.497-0.514]). PRS_{binary} and GAIL_{binary} identified more high-risk European than Asia individuals. High-risk proportions were higher among BrCa (16-26%) and DCIS (20-33%) compared to controls (9-15%) among young Europeans and all Asians. Fewer than 7% of BrCa, 10% of DCIS, and 3% of controls were classified as high-risk by multiple risk classifiers. Overlap between PRS_{binary} and PTV_{binary} alone (Europeans: 4.6%, Asians: 4.4%) and PRS_{binary} alone (Europeans: 13.9%, Asians: 8.5%). PRS_{binary} and FH_{binary} uniquely identified 5-6% and 9-11% of young BrCa, respectively.

Interpretation

The incomplete overlap between high-risk individuals identified by PRS_{binary} , $GAIL_{binary}$, FH_{binary} , and PTV_{binary} highlights the need for a comprehensive approach to breast cancer risk prediction.

SIGNIFICANCE

This study shows that different ways of predicting breast cancer risk do not always flag the same people, suggesting that combining multiple risk factors could improve early detection and screening.

KEYWORDS

Breast cancer, Ductal Carcinoma In Situ (DCIS), Polygenic risk score (PRS), Gail model, Risk stratification, BRCA1, BRCA2, risk-based screening

INTRODUCTION

A worldwide increase of 31% in the number of breast cancer cases is projected over the next two decades.¹ Early detection significantly improves survival rates.^{2,3} Multiple studies have shown that mammogram screenings reduce mortality rates for women above 50 years of age, while the benefits of screening for those younger are less clear.^{4,5} Current screening guidelines are based on age, yet many patients are diagnosed before reaching the recommended screening age.⁶

Advances in breast cancer research suggest the potential for more risk-based approaches to costeffective screening programs.⁷ Developed in the 1980s, the Gail model, a validated statistical tool, uses personal information to estimate breast cancer risk over the next five years. Originally developed for White females in the United States without a history of in situ or invasive breast cancer, its accuracy for non-White populations is debated.⁸⁻¹⁰ Common issues include both underestimation and overestimation in non-European populations, leading to unclear recommendations for diverse ethnic groups.^{9,11}

In addition to non-genetic risk factors, studies have explored the use of polygenic risk scores (PRS) to enhance existing prediction models.^{12,13} Breast cancer has a significant heritable component. While PRS has added value to prediction models, its implementation, particularly in Asian populations, remains inconclusive.¹⁵ This is partly because PRS training datasets have predominantly included European populations due to their larger representation in research.¹⁶

Protein-truncating variants (PTVs) are another genetic factor used in risk prediction. Unlike PRS, which aggregates the associated effects of numerous, relatively common genetic variants across the genome, PTVs specifically target variants that lead to premature protein termination, potentially disrupting gene function. This distinction means that PTVs are specific genetic changes with known functional impacts, while PRS provides an overview of associated genetic risks of small effect sizes across the genome. PTVs in nine breast cancer predisposition genes, *ATM, BRCA1, BRCA2, CHEK2, PALB2, BARD1, RAD51C, RAD51D,* or *TP53* have been shown clinically useful for inclusion on breast cancer risk prediction panels in a large analytical cohort comprising over 113,000 subjects and has been included in BOADICEA (Breast and Ovarian Analysis of Disease Incidence and Carrier Estimation Algorithm).^{17,18}

Studying the overlap of genetic and non-genetic risk factors in identifying high-risk individuals will provide us with information on complementary risk factors that will enhance our ability to identify the subgroup of the population who would benefit from risk reduction interventions. In this case-control analysis involving the Breast Cancer Consortium (BCAC) dataset, we explore how prediction tools—such as the Gail model, PRS, PTVs in known breast cancer predisposition genes, and family history—apply to both European and Asian populations across non-screening and screening age groups.

METHODS

Study population

BCAC is an international collaboration that was formed to provide large sample sizes for investigating genetic associations.¹⁹ Women diagnosed with invasive breast cancer (BrCa) or ductal carcinoma in situ patients (DCIS), and women with no prior diagnosis of breast cancer (controls) were recruited by study groups across the globe and collectively studied under BCAC.²⁰ Our study focuses on individuals who are genetically Asian or European-White (from here on referred to as "European").

To reduce the influence of missing values on the performance of the Gail model, studies with missing values for 50% or more for each of at least two of the three risk factors in the Gail model⁸ –age of menarche, age at first live birth, and first-degree family history of breast cancer–were excluded. The studies included are listed in **Supplementary Table 1**. Exclusion was done separately for individual studies and each disease status (BrCa, DCIS, and controls).

Further exclusions were made on an individual level (**Supplementary Figure 1**). Women with unknown age at enrolment for controls (n=5,566) and unknown age at diagnosis for BrCa and DCIS cases (n=2,103) were excluded. Women below the age of 30 years (n=2,360) and women above 80 years (n=1,897) for whom the Gail model prediction is not valid were excluded. A total of 180,398 individuals were included in our study. We compared demographic differences between the included and excluded individuals to assess potential selection bias. The result is presented in **Additional Material - Supplementary Table 2-**

Criteria to identify women at high risk of breast cancer

Four criteria were used to identify women at high risk of breast cancer: 1) 5-year absolute risk $\geq 1.66\%$ by the Gail model [GAIL_{binary}], 2) first-degree family history for breast cancer [FH_{binary}, yes/no], 3) 5-year absolute risk $\geq 1.66\%$ by a 313-variant breast cancer polygenic risk score¹⁴ [PRS_{binary}], and in a subset of women 4) carriers of pathogenic variants in breast cancer predisposition genes [PTV_{binary}]. The 1.66% five-year absolute risk threshold for breast cancer is widely adopted in clinical and research settings to reflect the level of risk at which women are considered for additional screening or preventive interventions, such as tamoxifen or raloxifene (from here onwards the high-risk category).²¹ Details of each risk factor are presented in **Additional Materials - Methods**.

Due to the large number of studies with varying degrees of missing data for different risk factors, the parsimonious Gail model, which most studies would have information on, was selected.⁸ The R package "BCRA" (version 2.1.2) was used to calculate 5-year absolute risk.⁸ Implementation is described in Additional Material-Genetic breast cancer risk factor. In our analysis, those with unknown family history were considered to have no family history.

We studied genetic risk based on common germline variants associated with breast cancer, using the breast cancer PRS with 313 variants calculated with PLINK (version 3) with the scoresum option.^{22,23} The 5-year absolute risk was calculated by estimating the theoretical odds ratio of this percentile in

relation to the 40-60 percentile, which is taken to represent the general population.²⁴ A subgroup of individuals (n_{European}=56,387, n_{Asian}=3,617) had both genotyping and targeted-sequencing data. Nine breast cancer predisposition genes (PTVs in *ATM, BRCA1, BRCA2, CHEK2, PALB2, BARD1, RAD51C, RAD51D*, or *TP53*) were studied collectively.

Statistical analysis

Differences in genetic and non-genetic breast cancer risk factors of BrCa, DCIS, and controls were assessed using the Chi-squared test (categorical variables) and Kruskal-Wallis test (continuous variables). Venn diagrams (R package "VennDiagram") were used to visualise the overlaps in high-risk individuals identified by the high-risk criteria. We subset the population by ancestry (European or Asian) and age (30-49 or 50-80 years). Overlaps between pairs of high-risk criteria were further considered by country. In the subset of individuals with both genotyping and targeted sequencing information (n=60,004), PRS_{binary} and PTV_{binary} (yes/no) were analysed for their ability to uniquely identify high-risk individuals.

Evaluating the drivers of Gail risk score in differentiating breast cancer cases from controls

Although studies with high missingness rates for variables required to compute the Gail risk score were excluded (see "Excluded participants" above), there were still individuals with missing values. Hence, we studied the potential drivers of the Gail risk score in discriminating BrCa cases from controls using logistic regression models. All combinations of risk factors, where available, were assessed. Discriminatory ability was assessed by the area under the receiver operator curve (AUC). Missing values were coded in accordance to the "BCRA".⁸

All analyses were performed in R version 4.2.2.

RESULTS

Cohort description

A total of 180,398 women were included, where 161,849 (90%) women were of European-ancestry and 18,549 (10%) were Asian-ancestry (**Table 1**). Of the European women, 68,540 (42%) were controls, and 83,685 (52%) were diagnosed with BrCa. Of the 18,549 Asian women, 8,347 (45%) were controls and 9,222 (50%) were BrCa cases (**Table 1**). In addition, there were 9,624 (6%) and 980 (5%) DCIS cases in European and Asian women, respectively (**Table 1**).

European-ancestry study population

The median age at diagnosis for BrCa cases of European ancestry was 57 years [interquartile range [IQR]: 49-65]. The corresponding age at interview for European-ancestry controls was 57 years [IQR: 50-64] (**Table 1**). BrCa cases were more likely to have a family history than controls (14% vs 9%, respectively). The distribution of the 5-year absolute risk by the Gail model was very similar between BrCa cases (median 1.25% [IQR: 0.89-1.73]) and controls (1.25% [IQR: 0.92-1.64]). The median 5-year absolute risk by PRS in BrCa cases was 0.95% [IQR: 0.62-1.44], significantly greater than that of

controls 0.69% [IQR: 0.46-1.05]. The PRS distributions (scoresum) and 5-year absolute risks were similar across countries (**Supplementary Figure 2**). Less than 40% of the population had PTV information. Out of 29,853 European BrCa patients, 1,927 (6%) were PTV carriers, which is three times the proportion found in the control group (583 out of 24,798, or 2%). The observations were largely similar for DCIS (**Table 1**).

Asian-ancestry study population

The median age at diagnosis for BrCa cases was 49 years [IQR: 43-57], and the age at enrolment was 50 years [IQR: 44-58] for controls (**Table 1**). Of the BrCa cases, 10% reported positive family history, while a smaller proportion of controls (6%) reported so. The distribution of 5-year absolute risk by the Gail model was not significantly different between BrCa cases and controls. The distribution of 5-year absolute risk by PRS for BrCa patients was shifted rightwards of controls (p<0.001). The distribution of PRS (scoresum) and 5-year absolute risks varied by country (**Supplementary Figure 3).** Among BrCa patients with known PTV information (n=2,178) 6% were mutation carriers. A smaller percentage (2%) of controls (n=1,115) were PTV carriers. As with the Europeans, the observations were mostly similar for DCIS in Asians (**Table 1**). and controls.

Associations between risk stratifiers, BrCa and DCIS

Table 2 and **Table 3** display the strengths of association of different risk stratifiers (PRS_{binary}, GAIL_{binary}, and FH_{binary}) with BrCa and DCIS, respectively, stratified by ancestry and age groups. Using both PRS_{binary} and GAIL_{binary} (i.e. individuals is stratified as high risk when either PRS or GAIL is $\geq 1.66\%$) improves the discriminatory ability as compared to using GAIL_{binary} alone (in Europeans: AUC_{BrCa-PRS_GAIL}=0.554 [0.552 to 0.557] vs AUC_{BrCa-GAIL}=0.522 [0.520-0.524]; in Asian: AUC_{BrCa-PRS_GAIL}=0.527 [0.523 to 0.532] vs AUC_{BrCa-GAIL}=0.506 [0.503-0.508]) (**Table 2**), In Europeans, the odds ratios and corresponding 95% confidence intervals for PRS_{binary} and GAIL_{binary} ($\geq 1.66\%$ 5-year absolute risk threshold) were 2.60 [2.52-2.69] and 1.25 [1.22-1.28], respectively, for BrCa, and 2.21 [2.08-2.35] and 1.21 [1.15-1.27], respectively, for DCIS. In Asians, PRS_{binary} showed significant associations with BrCa (1.63 [1.64-2.05]) and DCIS (2.30 [1.68-2.63]). The GAIL_{binary} showed associations with BrCa (1.59 [1.31-1.92]) and DCIS (2.13 [1.51-3.00]) in Asians. The effect sizes for the associations between PRS_{binary}, GAIL_{binary} and FH_{binary} were larger for younger Europeans than the older Europeans. In Asians, the same trend was observed for PRS_{binary} and BrCa, and FH_{binary} and DCIS.

Intersection of high-risk individuals identified by different risk factors

Figure 1 illustrates the overlap of high-risk individuals identified by PRS_{binary}, GAIL_{binary}, and FH_{binary} across different ancestry and age groups. For young Europeans and all Asians, the proportion of high-risk individuals among BrCa (16-26%) and DCIS (20-30%) cases was about twice that of the controls (9-13%). In these groups, women were primarily classified as high-risk due to FH_{binary} and PRS_{binary}. Less than 7%, 10%, and 3% of the BrCa, DCIS, and controls, respectively, were classified as high-risk by more than one criterion (i.e. PRS_{binary}, GAIL_{binary}, or FH_{binary}). PRS_{binary} uniquely identified 4-7%

of young BrCa and DCIS cases as high-risk, and 10-18% of older BrCa and DCIS cases as high-risk. FH_{binary} uniquely identified 9-11% of young BrCa and DCIS cases as high-risk, and 2-5% of older BrCa and DCIS cases as high-risk. The proportion of young Europeans (aged 30-49) uniquely identified by GAIL_{binary} to be at high-risk is 1%. Among Asians, all individuals classified as GAIL_{binary} high-risk were also positive for FH (i.e. 0% uniquely called by GAIL_{binary}). Among the older Europeans, 40% of the controls were identified as high-risk, compared to 52% of BrCa cases and 47% of DCIS cases.

Breast cancer predisposition genes (PTV_{binary}) and common variants (PRS_{binary}) identified different high-risk individuals

In the subgroup of individuals with target-enriched sequencing data, the proportion of women identified to be at high risk by both PRS_{binary} and PTV_{binary} was limited (0.6% of European and 0.1% of Asian) compared to PTV_{binary} alone (Europeans: 4.6%; Asians: 4.4%) and PRS_{binary} alone (Europeans: 13.9%; Asians: 8.5%) (**Table 4).** There were more older women than younger women at high risk due to PRS (2.3x in Europeans; 3.8x in Asians). Conversely, there were more younger women than older women at high risk due to PTV (1.9x in Europeans; 1.5x in Asians).

Country-specific differences in high-risk individuals identified by PRS and GAIL

Figure 2 shows the breakdown of BrCa/DCIS cases and controls identified to be at high risk by PRS_{binary} and GAIL_{binary} by country and age. Generally, both PRS_{binary} and GAIL_{binary} identified a higher proportion (%) of high-risk BrCa/DCIS individuals in the European-ancestry populations (median [IQR], PRS_{binary-young}: 5 [3-9], PRS_{binary-old}: 20 [10-23], GAIL_{binary-young}: 3 [2-7], GAIL_{binary-old}: 30 [24-39]) than the Asian countries (PRS_{binary-young}: 4 [2-5], PRS_{binary-old}: 15 [10-16], GAIL_{binary-young}: <1 [<1 to 2], GAIL_{binary-old}: 4 [2-10]).

Factors influencing the performance of the Gail model across different demographics

Figure 2 shows that both PRS_{binary} and GAIL_{binary} identified a higher proportion of high-risk individuals in European-ancestry compared to Asian-ancestry populations, with variations in performance by age. By studying the factors that influence the Gail model's performance across different demographics, we can better understand what drives the model's effectiveness in various populations. In **Figure 3A** and **Figure 3B**, we show that in Europeans, incorporating both family history (number of first-degree relatives with breast cancer) and prior breast biopsies was sufficient to achieve the highest AUC. For younger Asians, the key factors affecting model performance are age at menarche and the number of prior biopsies, with higher discriminatory ability observed in models that did not include family history (**Figure 3C**). In older Asians, the model's performance was not significantly different between those that included both age at first live birth and family history (**Figure 3D**). This indicates that, while the Gail model's performance for Europeans is primarily influenced by the number of first-degree relatives with breast cancer and prior biopsies, younger Asians are more influenced by age at first menarche, and for older Asians, both age at first live birth and family history are important.

Applying the high-risk criterion of \geq 1.66% 5-year absolute risk resulted in changes in the order of the models regarding their discriminatory ability (AUC) (**Supplementary Figure 4**). However, the AUCs of the models were small (highest AUC [95%CI]: 0.529 [0.526-0.531]) and not appreciably different (within 0.03 difference in AUCs. Excluding models with confidence intervals including 0.5, lowest AUC: 0.501 [0.500-0.501]) (**Supplementary Table 2**).

Discussion

We evaluated the performance of different risk stratifiers, including PRS, GAIL, and FH, in identifying high-risk individuals for BrCa and DCIS across various demographics and to understand the overlap and unique contributions of these models in different populations. The association between different risk stratifiers (PRS, GAIL, and FH) with BrCa and DCIS varies by ancestry and age. PRS demonstrated superior discrimination compared to the Gail model for predicting both BrCa and DCIS in European- and Asian-ancestry populations. Specifically, the 5-year absolute risk from PRS showed higher AUC values than the Gail model for both conditions. In Europeans, PRS and GAIL showed significant odds ratios for identifying high-risk individuals, with larger effect sizes observed in younger populations. In Asians, PRS and GAIL also showed significant associations. The overlap of high-risk individuals identified by PRS, GAIL, and FH revealed that PRS and FH were primary contributors to high-risk classification, particularly in young Europeans and all Asians. PRS uniquely identified a notable percentage of high-risk individuals that were missed by GAIL and FH, while GAIL identified a significant portion in older Europeans. Additionally, target-enriched sequencing data showed that high-risk individuals identified by both PRS and predisposition genes (PTV) were limited, with PRS alone identifying a larger proportion of high-risk individuals compared to PTV alone. Country-specific analysis indicated that both PRS and GAIL identified a higher proportion of high-risk individuals in European-ancestry populations compared to Asian-ancestry populations, with greater variability observed for GAIL.

Given the complexity and multifactorial nature of breast cancer, relying on a single risk factor or model may not sufficiently capture all high-risk individuals.^{6,12} The analysis of the intersection between high-risk individuals identified by PRS, GAIL, and FH reveals important insights into how these risk stratifiers overlap and uniquely contribute to risk assessment. Our results derived from the analysis of 180,398 woman across diverse ancestries corroborate previous findings that report a limited overlap in the high-risk individuals identified by different risk predictors.^{6,12} The unique contribution of PRS is particularly notable. Traditional risk models like the Gail model are less accurate in younger populations. These models often rely on risk factors which may not fully capture the risk in younger women. Younger women may not have a significant personal or reproductive history, making genetic information from PRS and specific gene mutations particularly valuable for risk assessment.^{6,12} No single model is suitable for every subgroup within the general population. The limited overlap and the unique contributions of each risk stratifier suggest that using a combination of these tools could provide a more comprehensive risk assessment, capturing high-risk individuals that might be missed by any single model. Comprehensive risk models such as BOADICEA improve prediction, however

they can be challenging to implement at the general population level.¹⁸ In addition, calibration and validation for populations not used in the model's development need to be done.

The evaluation of country-specific differences in high-risk identification by PRS and GAIL shows that European-ancestry populations generally had higher proportions of predicted high-risk individuals compared to Asian-ancestry populations. This is expected, as two risk predictors utilize breast cancer incidence rates that were higher in Europeans (i.e. the "White" used to develop the Gail model) than in Asians (i.e. the "Chinese").¹⁵ The variability in the performance of the GAIL model across different countries, with larger standard deviations compared to PRS, suggests that GAIL's effectiveness may be more influenced by regional factors, such as differences in reproductive factors, lifestyle and healthcare practices.^{11,25} The analysis of factors influencing the performance of the Gail model reveals differences in its effectiveness based on age and demographic factors. For Europeans, incorporating family history and prior breast biopsies achieved the highest AUC, emphasizing the importance of these factors in risk prediction. In younger Asians, age at menarche and the number of prior biopsies were more influential, with models excluding family history showing better performance. While there are likely regional differences in genetics, lifestyle, and healthcare practices, the Gail model may be compounded by variations in data quality and recall.^{25,26}

The current guidelines for breast cancer screening are based on sex and age. The recommendations typically advocate biennial mammography for women aged 50 to 69 or 70 years.^{3,27} Previously, the US Preventive Services Task Force (USPSTF) advised that the decision to begin biennial screening before age 50 should be personalized, considering the patient's values regarding specific benefits and harms. It is unclear if clinicians are provided with directives on the specific topics to discuss with patients regarding screening suitability. For age groups where the evidence for mammography is less definitive, integrating comprehensive risk stratification into discussions about screening would ensure that recommendations are as relevant as possible. However, in Apr 2024, the USPSTF published in its Final Recommendation Statement biennial mammogram screenings for all women aged 40 to 74 years to detect early-stage cancer (accessed Jul 23, 2024).²⁸ While this earlier screening will benefit many, it also raises concerns about over-screening and its potential consequences. Risk stratification could enhance the effectiveness of the new USPSTF guidelines by targeting screening efforts more precisely. By incorporating complementary individual risk factors, healthcare providers can better identify those who are genuinely at higher risk for breast cancer. As a result, risk stratification can help balance the benefits of early detection with the potential drawbacks of excessive screening.

Our study uses one of the largest breast cancer association study datasets, providing high statistical power for comprehensive risk factor analyses and diverse population coverage that includes both European and Asian ancestries. The study's multi-center scope allowed for the comparison of risk models across different countries. However, differences in study design, data collection methods, and risk factor definitions across included studies may have introduced variability and affected the consistency of results. Variations in the time of data collection and changes in clinical practices over time could affect the comparability of data across studies. Combining different studies introduced

gaps in data (i.e. missingness) for some risk factors, and exclusions of certain studies affected the generalizability of our findings for those regions. In addition, while the study covers women of European and Asian ancestries, it does not represent all global populations. Regional variations in absolute risk—due to genetic differences, varying gene predispositions, lifestyle factors, and healthcare access—also impact the applicability of risk models to other settings or demographics.^{11,29} Not all known breast cancer risk factors were considered in our analyses. Examples of other risk factors include mammographic density, physical activity, alcohol use, and smoking.^{18,30}

Overall, while PRS shows consistent superiority in breast cancer risk stratification across demographics, the complementary use of the Gail model and family history can enhance the overall risk assessment process. Integrating and calibrating these models for different ethnic populations, along with understanding their unique contributions and limitations, can lead to more precise identification of high-risk individuals who would otherwise be missed.¹³ Ideally, evaluating individual breast cancer risk could lead to more precise and cost-effective early detection by tailoring screening approaches to risk levels. However, despite advancements in risk assessment, it remains important to adhere to the established consensus guidelines for minimum mammography screening, as set by nationally recognized organizations with expertise in screening methodology.

DECLARATIONS Ethics approval

This study was approved by the A*STAR Institutional Review Board (reference number: 2022-041).

Consent to participate

Informed consent was obtained by the individual studies that contributed to BCAC.

Conflict of interest

The authors declare no potential conflicts of interest.

Availability of data and materials

The data used in our analyses are available upon reasonable request through BCAC, subject to data access committee approval.

Acknowledgements

This study is funded by the Agency for Science, Technology and Research (A*STAR) and PRECISION Health Research, Singapore (PRECISE). The breast cancer genome-wide association analyses in BCAC were supported by the Government of Canada through Genome Canada and the Canadian Institutes of Health Research, the 'Ministère de l'Économie, de la Science et de l'Innovation du Québec' through Genome Québec and grant PSR-SIIRI-701, The National Institutes of Health (U19 CA148065, X01HG007492), Cancer Research UK (C1287/A10118, C1287/A16563, C1287/A10710), and The European Union (HEALTH-F2-2009-223175 and H2020 633784 and 634935). All studies and funders are listed in Additional Materials (BCAC Funding and Acknowledgments).

Author contributions

Conception and design of the study: JL; Acquisition of data: All; Analysis and interpretation of data: PJH, MHG, CKYL, JL; Drafting of the manuscript or revising it for important content: JL, PJH, CKYL; Final approval of the version submitted for publication: All.

Table 1. Characteristics of 161,849 European-ancestry and 18,549 Asian-ancestry individuals (non-breast cancer controls (controls), patients diagnosed with invasive breast cancer (BrCa) and patients diagnosed with DCIS (DCIS)) between ages 30 and 80 years. IQR interquartile range. Family history: number of first-degree relatives with breast cancer.

_	European-ancestry, n=161,849			Asian-ancestry, n=18,549			
	Controls, n=68,540 (42%)	BrCa, n=83,685 (52%)	DCIS, n=9,624 (6%)	Controls, n= 8,347 (45%)	BrCa, n=9,222 (50%)	DCIS, n=980 (5%)	
Median age (at interview/ diagnosis), years (IQR)	57 (50 to 64)	57 (49 to 65)	55 (50 to 63)	50 (44 to 58)	49 (43 to 57)	49 (43 to 56)	
Age at menarche, n (%)							
<12 years	9,825 (14)	10,859 (13)	1,733 (18)	508 (6)	586 (6)	61 (6)	
12 to 14 years	30,243 (44)	33,476 (40)	4,546 (47)	3,080 (37)	3,278 (36)	326 (33)	
>=14 years	21,133 (31)	23,169 (28)	2,758 (29)	4,321 (52)	4,186 (45)	467 (48)	
Unknown	7,339 (11)	16,181 (19)	587 (6)	438 (5)	1,172 (13)	126 (13)	
Age at first full-term pregnancy, n (%)							
Nulliparous	8,319 (12)	10,769 (13)	1,520 (16)	1,005 (12)	1,236 (13)	158 (16)	
<20 years	5,427 (8)	6,138 (7)	794 (8)	316 (4)	362 (4)	29 (3)	
20 to 24 years	21,863 (32)	21,927 (26)	2,899 (30)	1,948 (23)	1,810 (20)	135 (14)	
25 to 29 years	17,543 (26)	18,174 (22)	2,382 (25)	2,836 (34)	3,030 (33)	293 (30)	
>=30 years	8,487 (12)	9,973 (12)	1,386 (14)	1,149 (14)	1,554 (17)	156 (16)	
Unknown	6,901 (10)	16,704 (20)	643 (7)	1,093 (13)	1,230 (13)	209 (21)	
Family history, n (%)							
No	45,629 (67)	48,903 (58)	4,005 (42)	7,181 (86)	7,500 (81)	708 (72)	
1	5,348 (8)	10,256 (12)	1,115 (12)	437 (5)	876 (9)	114 (12)	
2+	791 (1)	2,064 (2)	305 (3)	63 (1)	93 (1)	13 (1)	
Unknown	16,772 (24)	22,462 (27)	4,199 (44)	666 (8)	753 (8)	145 (15)	
Number of breast biopsy, n (%)							
No	930 (1)	3,181 (4)	96 (1)	0 (0)	0 (0)	0 (0)	
1	277 (0)	3,148 (4)	216 (2)	0 (0)	0 (0)	0 (0)	
2+	103 (0)	1,822 (2)	148 (2)	0 (0)	0 (0)	0 (0)	
Unknown	67,230 (98)	75,534 (90)	9,164 (95)	8,347 (100)	9,222 (100)	980 (100)	
Atypical hyperplasia, n (%)							
No	930 (1)	3,181 (4)	96 (1)	0 (0)	0 (0)	0 (0)	
Yes	5 (0)	49 (0)	8 (0)	0 (0)	0 (0)	0 (0)	
Unknown	67,605 (99)	80,455 (96)	9,520 (99)	8,347 (100)	9,222 (100)	980 (100)	
Median 5-year absolute risk by Gail (IQR)	1.25 (0.92 to 1.64)	1.25 (0.89 to 1.73)	1.29 (0.98 to 1.71)	0.61 (0.45 to 0.78)	0.61 (0.44 to 0.80)	0.61 (0.38 to 0.81)	
Protein truncating variants (9 Genes)							
No	24,215 (35)	27,926 (33)	1,662 (17)	1,091 (13)	2,049 (22)	317 (32)	
Yes	583 (1)	1,927 (2)	74 (1)	24 (0)	129 (1)	7 (1)	
Unknown	43,742 (64)	53,832 (64)	7,888 (82)	7,232 (87)	7,044 (76)	656 (67)	
Polygenic risk score (PRS)	-0.45 (-0.86 to -0.04)	-0.09 (-0.51 to 0.32)	-0.15 (-0.55 to 0.27)	0·16 (-0·20 to 0·53)	0.37 (-0.01 to 0.76)	0.45 (0.05 to 0.83)	
Median 5-year absolute risk by PRS (IQR)	0.69 (0.46 to 1.05)	0.95 (0.62 to 1.44)	0.91 (0.61 to 1.37)	0.62 (0.40 to 0.95)	0.74 (0.44 to 1.15)	0.79 (0.44 to 1.28)	

16

Table 2. Association between high-risk criteria and case-control status (invasive breast cancer cases/ non-breast cancer controls), using univariate logistic regression. Analysis was repeated by age categories 30 to 49 years ($n_{European}=40,306$, $n_{Asian}=8,456$) and 50 to 80 years ($n_{European}=111,919$, $n_{Asian}=9,113$). PRS: 5-year absolute risk using polygenic risk score $\geq 1.66\%$. GAIL: 5-year absolute risk using the Gail model $\geq 1.66\%$. FH: having at least one first-degree family history of breast cancer. * High: individuals who were identified by any of the criteria were classified as "Yes". OR: odds ratio, CI: confidence interval, P: p-value. 5-yr absolute risk: 5-year absolute risk (continuous).

		All ages		Age 30 to 49 years			Age 50 to 80 years		
	OR (95% CI)	P	AUC (95% CI)	OR (95% CI)	Р	AUC (95% CI)	OR (95% CI)	Р	AUC (95% CI)
European-ancestry, n=152,225			· · ·			<u>, , , , , , , , , , , , , , , , , </u>			
5-yr abs risk by PRS									
(continuous)	1.97 (1.94 to 2.01)	<0.001	0.635 (0.632 to 0.638)	2.51 (2.39 to 2.62)	<0.001	0.622 (0.617 to 0.628)	2.06 (2.02 to 2.11)	<0.001	0.653 (0.650 to 0.656)
5-yr abs risk by Gail	, , , , , , , , , , , , , , , , , , ,			,			, , , , , , , , , , , , , , , , , , ,		
(continuous)	1.12 (1.11 to 1.14)	<0.001	0.492 (0.489 to 0.495)	1.35 (1.29 to 1.40)	<0.001	0.493 (0.487 to 0.499)	1.18 (1.16 to 1.19)	<0.001	0.517 (0.514 to 0.520)
PRS (Yes, ref=No)	2.60 (2.52 to 2.69)	<0.001	0.552 (0.550 to 0.554)	3.35 (3.02 to 3.70)	<0.001	0.530 (0.528 to 0.532)	2.64 (2.55 to 2.74)	<0.001	0.562 (0.560 to 0.564)
Gail (Yes, ref=No)	1.25 (1.22 to 1.28)	<0.001	0.522 (0.520 to 0.524)	2.61 (2.36 to 2.88)	<0.001	0.523 (0.521 to 0.526)	1.27 (1.24 to 1.30)	<0.001	0.527 (0.524 to 0.529)
FH (Yes, ref=No) PRS/GAIL*	1.56 (1.51 to 1.60)	<0.001	0.529 (0.527 to 0.530)	1.87 (1.76 to 1.98)	<0.001	0.538 (0.535 to 0.542)	1.47 (1.42 to 1.52)	<0.001	0.525 (0.523 to 0.527)
(Yes, ref=No) PRS/GAIL/FH* (Yes,	1.62 (1.58 to 1.65)	<0.001	0.554 (0.552 to 0.557)	2.98 (2.77 to 3.21)	<0.001	0.548 (0.545 to 0.551)	1.67 (1.63 to 1.71)	<0.001	0.563 (0.560 to 0.565)
ref=No)	1.63 (1.60 to 1.67)	<0.001	0.558 (0.556 to 0.560)	2·34 (2·22 to 2·47)	<0.001	0.566 (0.562 to 0.569)	1.63 (1.59 to 1.67)	<0.001	0.560 (0.557 to 0.563)
Asian-ancestry, n=17 569									
5-vr abs risk by PRS									
(continuous)	1.48 (1.41 to 1.56)	<0.001	0.564 (0.556 to 0.573)	1.62 (1.47 to 1.78)	<0.001	0.551 (0.539 to 0.563)	1.64 (1.53 to 1.75)	<0.001	0.600 (0.588 to 0.611)
5-vr abs risk by Gail		10 001		1 02 (1 11 10 1 10)	10 001			10 001	
(continuous)	1.19 (1.09 to 1.30)	<0.001	0.506 (0.497 to 0.514)	0.94 (0.81 to 1.08)	0.377	0.523 (0.511 to 0.535)	1.82 (1.61 to 2.07)	<0.001	0.554 (0.543 to 0.566)
PRS (Yes, ref=No)	1.83 (1.64 to 2.05)	<0.001	0.523 (0.519 to 0.527)	2.24 (1.76 to 2.85)	<0.001	0.515 (0.511 to 0.519)	1.89 (1.66 to 2.14)	<0.001	0.535 (0.528 to 0.542)
Gail (Yes. ref=No)	1.59 (1.31 to 1.92)	<0.001	0.506 (0.503 to 0.508)	1.34 (0.84 to 2.13)	0.217	0.501 (0.499 to 0.503)	1.77 (1.44 to 2.19)	<0.001	0.512 (0.507 to 0.516)
FH (Yes, ref=No) PRS/GAIL*	1.97 (1.76 to 2.19)	<0.001	0.527 (0.523 to 0.531)	1.87 (1.59 to 2.19)	<0.001	0.524 (0.518 to 0.530)	2.07 (1.78 to 2.41)	<0.001	0.530 (0.524 to 0.535)
(Yes, ref=No) PRS/GAIL/FH* (Yes,	1.78 (1.61 to 1.96)	<0.001	0.527 (0.523 to 0.532)	2.02 (1.63 to 2.50)	<0.001	0.516 (0.511 to 0.520)	1.89 (1.69 to 2.12)	<0.001	0.543 (0.536 to 0.551)
ref=No)	1.91 (1.76 to 2.08)	<0.001	0.544 (0.539 to 0.550)	1.95 (1.70 to 2.24)	<0.001	0.535 (0.528 to 0.542)	2.01 (1.81 to 2.23)	<0.001	0.556 (0.548 to 0.564)

Table 3. Association between high-risk criteria and case-control status (DCIS cases/ non-breast cancer controls), using univariate logistic regression. Analysis was repeated by age categories 30 to 49 years ($n_{European}=19,424$, $n_{Asian}=4,322$) and 50 to 80 years ($n_{European}=58,740$, $n_{Asian}=5,005$). PRS: 5-year absolute risk using polygenic risk score $\geq 1.66\%$. GAIL: 5-year absolute risk using the Gail model $\geq 1.66\%$. FH: having at least one first-degree family history of breast cancer. * High: individuals who were identified by any of the three criteria (PRS, GAIL, FH) were classified as "Yes". OR: odds ratio, CI: confidence interval, P: p-value. 5-yr absolute risk (continuous).

	All ages		Age 30 to 49 years			Age 50 to 80 years			
	OR (95% CI)	P	AUC (95% CI)	OR (95% CI)	Р	AUC (95% CI)	OR (95% CI)	Р	AUC (95% CI)
European-ancestry,									· · · ·
n=78,164									
5-yr abs risk by PRS (continuous)	1.63 (1.59 to 1.68)	<0.001	0.626 (0.620 to 0.631)	2.56 (2.37 to 2.78)	<0.001	0.657 (0.645 to 0.669)	1.56 (1.51 to 1.61)	<0.001	0.620 (0.613 to 0.626)
5-yr abs risk by Gail (continuous)	1.23 (1.20 to 1.26)	<0.001	0.537 (0.531 to 0.543)	2.28 (2.10 to 2.49)	<0.001	0.610 (0.597 to 0.622)	1.19 (1.16 to 1.22)	<0.001	0.519 (0.512 to 0.526)
PRS (Yes, ref=No)	2.21 (2.08 to 2.35)	<0.001	0.541 (0.537 to 0.544)	3.30 (2.78 to 3.91)	<0.001	0.529 (0.523 to 0.535)	2.11 (1.98 to 2.26)	<0.001	0.544 (0.539 to 0.549)
Gail (Yes, ref=No)	1.21 (1.15 to 1.27)	<0.001	0.518 (0.513 to 0.523)	3.35 (2.85 to 3.94)	<0.001	0.534 (0.527 to 0.540)	1.12 (1.06 to 1.18)	<0.001	0.512 (0.506 to 0.518)
FH (Yes, ref=No) PRS/GAIL*	1.75 (1.65 to 1.85)	<0.001	0.537 (0.533 to 0.542)	2.24 (2.00 to 2.50)	<0.001	0.553 (0.544 to 0.561)	1.62 (1.52 to 1.72)	<0.001	0.532 (0.528 to 0.537)
(Yes, ref=No)	1.46 (1.40 to 1.53)	<0.001	0.542 (0.537 to 0.547)	3.50 (3.09 to 3.96)	< 0.001	0.559 (0.551 to 0.567)	1.34 (1.28 to 1.41)	<0.001	0.536 (0.530 to 0.542)
PRS/GAIL/FH* (Yes, ref=No)	1.52 (1.45 to 1.59)	<0.001	0.549 (0.544 to 0.554)	2.74 (2.48 to 3.02)	<0.001	0.581 (0.572 to 0.591)	1.36 (1.29 to 1.43)	<0.001	0.538 (0.531 to 0.544)
Asian-ancestry,									
5-yr abs risk by PRS (continuous)	1.67 (1.52 to 1.83)	<0.001	0.587 (0.566 to 0.607)	1.70 (1.42 to 2.03)	<0.001	0.556 (0.528 to 0.584)	1.89 (1.69 to 2.12)	<0.001	0.654 (0.628 to 0.680)
5-yr abs risk by Gail (continuous)	1.25 (1.03 to 1.52)	0.023	0.507 (0.486 to 0.528)	0.99 (0.71 to 1.38)	0.93	0.533 (0.505 to 0.562)	1.88 (1.46 to 2.41)	<0.001	0.542 (0.513 to 0.572)
PRS (Yes, ref=No)	2.30 (1.88 to 2.83)	<0.001	0.535 (0.524 to 0.546)	2.22 (1.43 to 3.43)	<0.001	0.514 (0.504 to 0.525)	2.65 (2.09 to 3.36)	<0.001	0.561 (0.542 to 0.580)
Gail (Yes, ref=No)	2.13 (1.51 to 3.00)	<0.001	0.511 (0.505 to 0.518)	1.86 (0.81 to 4.26)	0.144	0.503 (0.498 to 0.509)	2.41 (1.64 to 3.53)	<0.001	0.521 (0.509 to 0.533)
FH (Yes, ref=No) PRS/GAIL*	2.76 (2.28 to 3.35)	<0.001	0.547 (0.535 to 0.558)	2.89 (2.20 to 3.78)	<0.001	0.550 (0.533 to 0.566)	2.64 (2.00 to 3.48)	<0.001	0.544 (0.527 to 0.560)
(Yes, ref=No)	2.23 (1.85 to 2.69)	<0.001	0.542 (0.530 to 0.554)	2.09 (1.41 to 3.10)	< 0.001	0.517 (0.505 to 0.528)	2.63 (2.11 to 3.28)	<0.001	0.573 (0.553 to 0.593)
PRS/GAIL/FH* (Yes, ref=No)	2.58 (2.21 to 3.02)	<0.001	0.571 (0.557 to 0.585)	2.63 (2.06 to 3.37)	<0.001	0.557 (0.539 to 0.575)	2.76 (2.24 to 3.39)	<0.001	0.589 (0.567 to 0.611)

Table 4. Overlap of individuals with 5-year absolute risk by polygenic risk score ≥1.66% (PRS) and carriers of protein-truncating variants in at least one of nine breast cancer predisposition genes (i.e. PTVs in *ATM, BRCA1, BRCA2, CHEK2, PALB2, BARD1, RAD51C, RAD51D,* or *TP53*).

			-		High-risk	
Status	Ancestry	Age, years	n	PRS	PTV	PRS + PTV
All	European	30 to 80	56,387	7,844 (13-91)	2,584 (4.58)	350 (0.62)
		30 to 49	13,843	971 (7.01)	1,002 (7-24)	71 (0.51)
		50 to 80	42,544	6,873 (16.16)	1,582 (3.72)	279 (0.66)
	Asian	30 to 80	3,617	306 (8.46)	160 (4-42)	5 (0.14)
		30 to 49	2,358	101 (4.28)	118 (5.00)	3 (0.13)
		50 to 80	1,259	205 (16-28)	42 (3.34)	2 (0.16)
Control	European	30 to 80	24,798	2,022 (8-15)	583 (2.35)	39 (0.16)
		30 to 49	4,860	136 (2.80)	173 (3.56)	3 (0.06)
		50 to 80	19,938	1,886 (9.46)	410 (2.06)	36 (0.18)
	Asian	30 to 80	1,115	58 (5.20)	24 (2-15)	1 (0.09)
		30 to 49	769	21 (2.73)	18 (2-34)	1 (0.13)
		50 to 80	346	37 (10.69)	6 (1.73)	0 (0.00)
Invasive breast	European	30 to 80	29,853	5,515 (18-47)	1,927 (6·45)	307 (1.03)
cancer		30 to 49	8,556	789 (9-22)	808 (9-44)	67 (0.78)
		50 to 80	21,297	4,726 (22.19)	1,119 (5-25)	240 (1.13)
	Asian	30 to 80	2,178	208 (9.55)	129 (5.92)	4 (0.18)
		30 to 49	1,400	70 (5.00)	95 (6-79)	2 (0.14)
		50 to 80	778	138 (17.74)	34 (4.37)	2 (0.26)
DCIS	European	30 to 80	1,736	307 (17.68)	74 (4-26)	4 (0-23)
		30 to 49	427	46 (10.77)	21 (4-92)	1 (0-23)
		50 to 80	1,309	261 (19-94)	53 (4.05)	3 (0-23)
	Asian	30 to 80	324	40 (12·35)	7 (2.16)	0 (0.00)
		30 to 49	189	10 (5·29)	5 (2.65)	0 (0.00)
		50 to 80	135	30 (22.22)	2 (1.48)	0 (0.00)

Figure 1. Venn diagram depicting the overlaps between individuals identified as high risk by the three criteria. Breast cancer polygenic risk score (PRS, 5-year absolute risk using polygenic risk score ≥1.66%), the Gail model (GAIL, 5-year absolute risk using the Gail model ≥1.66%), and family history (FH, having at least one first-degree family history of breast cancer). European: women of European ancestry; Asian: women of Asian ancestry.



FH High risk, n=691 (30%) Total, n=2342

PBS GAIL <0.1 0 Controls 8

FH High risk, n=2267 (13%) Total, n=17082



(B) European, 50 to 80 years

FF High risk, n=3441 (47%) Total, n=7282



Total, n=51458



(C) Asian, 30 to 49 years

High risk, n=731 (16%) Total, n=4634



High risk, n=101 (20%) Total, n=500



High risk, n=335 (9%) Total. n=3822



High risk, n=1208 (26%) Total, n=4588

(D) Asian, 50 to 80 years



High risk, n=158 (33%) Total, n=480



High risk, n=684 (15%) Total, n=4525

Figure 2. The proportion of individuals identified as at high risk by the breast cancer polygenic risk score (PRS) and the Gail model (GAIL), by country and age. The proportion of individuals identified as high risk by both criteria is indicated in grey. PRS: 5-year absolute risk using polygenic risk score ≥1.66%. GAIL: 5-year absolute risk using the Gail model ≥1.66%. Numbers adjacent to the bars represent the number of high-risk individuals identified by respective risk tools. European: women of European ancestry; Asian: women of Asian ancestry.





Eu	ropean			
	PR	S	G	AIL
	Control	BrCa/DC	IS Contro	BrCa/DCIS
Australia	298	620	829	1068
Belarus		68		18
Belgium	- 1	465		266
Canada	118	379	509	928
Denmark	1	11		24
Finland	22	98	40	150
France	82	249	209	368
Germany	415	2323	1205	3110
Greece		120		107
Israel	66	267	157	371
Italy	50	214	84	300
Netherlands	109	416	268	746
Poland	175	317	365	500
Spain	61	303	137	435
Sweden	1150	2388	4498	5072
UK	906	2918	2770	3366
USA	1533	3544	4850	7510

	Asian		
	Control	BrCa/DCIS	Control BrCa/DCIS
Canada	12	35	8 24
China	117	173	9 31
Hong Kong	20	20	17 23
Japan	137	146	30 27
Korea	29	77	8 37
Malaysia	76	104	59 56
Singapore	- 11	96	o
Taiwan	19	130	7 19
Thailand	17	44	38
USA	1	34	2 61
	0.30.10 Proport	0.10.30.5 tion identi	0.30.100.10.30.5

Criterion PRS only Gail only Both

10

18

01 0

Proportion identified as high risk

Korea 11 52

Malaysia 14 26

Singapore Taiwan

Thailand

USA

Figure 3. Discriminatory ability of risk factor combinations in the Gail model. The five-year absolute risk was calculated using the R package "BCRA" and used to predict the invasive breast cancer case-control status of the individuals. Dots represent risk factors included in the model, and crosses indicate the model with all risk factors with the addition of atypical hyperplasia. European: women of European ancestry; Asian: women of Asian ancestry.



References

- 1 Sedeta, E. T., Jobre, B. & Avezbakiyev, B. Breast cancer: Global patterns of incidence, mortality, and trends. *Journal of Clinical Oncology* **41**, 10528-10528, doi:10.1200/JCO.2023.41.16_suppl.10528 (2023).
- 2 Ginsburg, O., Yip, C. H., Brooks, A., Cabanes, A., Caleffi, M., Dunstan Yataco, J. A. *et al.* Breast cancer early detection: A phased approach to implementation. *Cancer* **126 Suppl 10**, 2379-2393, doi:10.1002/cncr.32887 (2020).
- Lim, Y. X., Lim, Z. L., Ho, P. J. & Li, J. Breast Cancer in Asia: Incidence, Mortality, Early Detection, Mammography Programs, and Risk-Based Screening Initiatives. *Cancers (Basel)* **14**, doi:10.3390/cancers14174218 (2022).
- 4 Lim, Z. L., Ho, P. J., Khng, A. J., Yeoh, Y. S., Ong, A. T. W., Tan, B. K. T. *et al.* Mammography screening is associated with more favorable breast cancer tumor characteristics and better overall survival: case-only analysis of 3739 Asian breast cancer patients. *BMC Med* **20**, 239, doi:10.1186/s12916-022-02440-y (2022).
- 5 Evans, A. & Whelehan, P. Breast screening policy: are we heading in the right direction? *Clin Radiol* 66, 915-919, doi:10.1016/j.crad.2011.03.024 (2011).
- 6 Ho, P. J., Ho, W. K., Khng, A. J., Yeoh, Y. S., Tan, B. K., Tan, E. Y. *et al.* Overlap of high-risk individuals predicted by family history, and genetic and non-genetic breast cancer risk prediction models: implications for risk stratification. *BMC Med* **20**, 150, doi:10.1186/s12916-022-02334-z (2022).
- 7 Sun, L., Legood, R., Sadique, Z., Dos-Santos-Silva, I. & Yang, L. Cost-effectiveness of risk-based breast cancer screening program, China. *Bull World Health Organ* 96, 568-577, doi:10.2471/BLT.18.207944 (2018).
- 8 Gail, M. H., Brinton, L. A., Byar, D. P., Corle, D. K., Green, S. B., Schairer, C. *et al.* Projecting individualized probabilities of developing breast cancer for white females who are being examined annually. *J Natl Cancer Inst* 81, 1879-1886, doi:10.1093/jnci/81.24.1879 (1989).
- 9 Ho, P. J., Wong, F. Y., Chay, W. Y., Lim, E. H., Lim, Z. L., Chia, K. S. *et al.* Breast cancer risk stratification for mammographic screening: A nationwide screening cohort of 24,431 women in Singapore. *Cancer Med* **10**, 8182-8191, doi:10.1002/cam4.4297 (2021).
- 10 Mertens, E., Barrenechea-Pulache, A., Sagastume, D., Vasquez, M. S., Vandevijvere, S. & Penalvo, J. L. Understanding the contribution of lifestyle in breast cancer risk prediction: a systematic review of models applicable to Europe. *BMC Cancer* 23, 687, doi:10.1186/s12885-023-11174-w (2023).
- 11 Stevanato, K. P., Pedroso, R. B., Dell Agnolo, C. M., Santos, L. D., Pelloso, F. C., Carvalho, M. D. B. *et al.* Use and Applicability of the Gail Model to Calculate Breast Cancer Risk: A Scoping Review. *Asian Pac J Cancer Prev* **23**, 1117-1123, doi:10.31557/APJCP.2022.23.4.1117 (2022).
- 12 Ho, P. J., Lim, E. H., Hartman, M., Wong, F. Y. & Li, J. Breast cancer risk stratification using genetic and non-genetic risk assessment tools for 246,142 women in the UK Biobank. *Genet Med* **25**, 100917, doi:10.1016/j.gim.2023.100917 (2023).
- 13 Li, S. X., Milne, R. L., Nguyen-Dumont, T., Wang, X., English, D. R., Giles, G. G. et al. Prospective Evaluation of the Addition of Polygenic Risk Scores to Breast Cancer Risk Models. JNCI Cancer Spectr 5, doi:10.1093/jncics/pkab021 (2021).
- 14 Mavaddat, N., Michailidou, K., Dennis, J., Lush, M., Fachal, L., Lee, A. *et al.* Polygenic Risk Scores for Prediction of Breast Cancer and Breast Cancer Subtypes. *Am J Hum Genet* **104**, 21-34, doi:10.1016/j.ajhg.2018.11.002 (2019).
- 15 Ho, W. K., Tai, M. C., Dennis, J., Shu, X., Li, J., Ho, P. J. *et al.* Polygenic risk scores for prediction of breast cancer risk in Asian populations. *Genet Med* **24**, 586-600, doi:10.1016/j.gim.2021.11.008 (2022).
- 16 Martin, A. R., Kanai, M., Kamatani, Y., Okada, Y., Neale, B. M. & Daly, M. J. Clinical use of current polygenic risk scores may exacerbate health disparities. *Nat Genet* **51**, 584-591, doi:10.1038/s41588-019-0379-x (2019).

- 17 Breast Cancer Association, C., Dorling, L., Carvalho, S., Allen, J., Gonzalez-Neira, A., Luccarini, C. *et al.* Breast Cancer Risk Genes - Association Analysis in More than 113,000 Women. *N Engl J Med* **384**, 428-439, doi:10.1056/NEJMoa1913948 (2021).
- 18 Lee, A., Mavaddat, N., Cunningham, A., Carver, T., Ficorella, L., Archer, S. *et al.* Enhancing the BOADICEA cancer risk prediction model to incorporate new data on RAD51C, RAD51D, BARD1 updates to tumour pathology and cancer incidence. *J Med Genet* **59**, 1206-1218, doi:10.1136/jmedgenet-2022-108471 (2022).
- 19 Breast Cancer Association, C. Commonly studied single-nucleotide polymorphisms and breast cancer: results from the Breast Cancer Association Consortium. *J Natl Cancer Inst* **98**, 1382-1396, doi:10.1093/jnci/djj374 (2006).
- 20 Petridis, C., Brook, M. N., Shah, V., Kohut, K., Gorman, P., Caneppele, M. *et al.* Genetic predisposition to ductal carcinoma in situ of the breast. *Breast Cancer Res* 18, 22, doi:10.1186/s13058-016-0675-7 (2016).
- 21 Smith, S. G., Sestak, I., Forster, A., Partridge, A., Side, L., Wolf, M. S. *et al.* Factors affecting uptake and adherence to breast cancer chemoprevention: a systematic review and meta-analysis. *Ann Oncol* **27**, 575-590, doi:10.1093/annonc/mdv590 (2016).
- Purcell, S., Neale, B., Todd-Brown, K., Thomas, L., Ferreira, M. A., Bender, D. *et al.* PLINK: a tool set for whole-genome association and population-based linkage analyses. *Am J Hum Genet* **81**, 559-575, doi:10.1086/519795 (2007).
- 23 Mavaddat, N., Pharoah, P. D., Michailidou, K., Tyrer, J., Brook, M. N., Bolla, M. K. *et al.* Prediction of breast cancer risk based on profiling with common genetic variants. *J Natl Cancer Inst* **107**, doi:10.1093/jnci/djv036 (2015).
- 24 Wen, W., Shu, X. O., Guo, X., Cai, Q., Long, J., Bolla, M. K. *et al.* Prediction of breast cancer risk based on common genetic variants in women of East Asian ancestry. *Breast Cancer Res* **18**, 124, doi:10.1186/s13058-016-0786-1 (2016).
- 25 Pruitt, W. R., Samuels, B. & Cunningham, S. The Gail Model and Its Use in Preventive Screening: A Comparison of the Corbelli Study. *Cureus* **16**, e56290, doi:10.7759/cureus.56290 (2024).
- 26 Coughlin, S. S. Recall bias in epidemiologic studies. *J Clin Epidemiol* **43**, 87-91, doi:10.1016/0895-4356(90)90060-3 (1990).
- 27 Ebell, M. H., Thai, T. N. & Royalty, K. J. Cancer screening recommendations: an international comparison of high income countries. *Public Health Rev* **39**, 7, doi:10.1186/s40985-018-0080-0 (2018).
- 28 Final Recommendation Statement Breast Cancer: Screening, <<u>https://www.uspreventiveservicestaskforce.org/uspstf/recommendation/breastcancer-screening</u>> (2024).
- 29 Yiangou, K., Mavaddat, N., Dennis, J., Zanti, M., Wang, Q., Bolla, M. K. *et al.* Differences in polygenic score distributions in European ancestry populations: implications for breast cancer risk prediction. *medRxiv*, doi:10.1101/2024.02.12.24302043 (2024).
- Pettersson, A., Graff, R. E., Ursin, G., Santos Silva, I. D., McCormack, V., Baglietto, L. *et al.* Mammographic density phenotypes and risk of breast cancer: a meta-analysis. *J Natl Cancer Inst* **106**, doi:10.1093/jnci/dju078 (2014).