Title

Polygenic score distribution differences across European ancestry populations: implications for breast cancer risk prediction.

Journal

Breast Cancer Research, 26(1)

Authors

Yiangou, Kristia

Mavaddat, Nasim

Dennis, Joe

et al.

Publication Date

2024-12-29

DOI

10.1186/s13058-024-01947-x

Peer reviewed

Breast Cancer Research

RESEARCH

Open Access

Polygenic score distribution differences across European ancestry populations: implications for breast cancer risk prediction

Kristia Yiangou[1], Nasim Mavaddat[2], Joe Dennis[2], Maria Zanti[1], Qin Wang[2], Manjeet K. Bolla[2], Mustapha Abubakar[3], Thomas U. Ahearn[3], Irene L. Andrulis[4,5], Hoda Anton-Culver[6], Natalia N. Antonenkova[7], Volker Arndt[8], Kristan J. Aronson[9], Annelie Augustinsson[10], Adinda Baten[11], Sabine Behrens[12], Marina Bermisheva[13,14], Amy Berrington de Gonzalez[15], Katarzyna Białkowska[16], Nicholas Boddicker[17], Clara Bodelon[18], Natalia V. Bogdanova[7,19,20], Stig E. Bojesen[21,22,23], Kristen D. Brantley[24], Hiltrud Brauch[25,26,27], Hermann Brenner[8,28], Nicola J. Camp[29], Federico Canzian[30], Jose E. Castelao[31], Melissa H. Cessna[32,33], Jenny Chang-Claude[12,34], Georgia Chenevix-Trench[35], Wendy K. Chung[36], NBCS Collaborators[37,38,39,40,41,42,43,44,45,46,47,48], Sarah V. Colonna[29], Fergus J. Couch[49], Angela Cox[50], Simon S. Cross[51], Kamila Czene[52], Mary B. Daly[53], Peter Devilee[54,55], Thilo Dörk[20], Alison M. Dunning[56], Diana M. Eccles[57], A. Heather Eliassen[24,58,59], Christoph Engel[60,61], Mikael Eriksson[52], D. Gareth Evans[62,63], Peter A. Fasching[64], Olivia Fletcher[65], Henrik Flyger[66], Lin Fritschi[67], Manuela Gago-Dominguez[68], Aleksandra Gentry-Maharaj[69,70], Anna González-Neira[71,72], Pascal Guénel[73], Eric Hahnen[74,75], Christopher A. Haiman[76], Ute Hamann[77], Jaana M. Hartikainen[78,79], Vikki Ho[80], James Hodge[18], Antoinette Hollestelle[81], Ellen Honisch[82], Maartje J. Hooning[81], Reiner Hoppe[25,83], John L. Hopper[84], Sacha Howell[85,86,87], Anthony Howell[88], ABCTB Investigators[89], kConFab Investigators[90,91], Simona Jakovchevska[92], Anna Jakubowska[16,93], Helena Jernström[10], Nichola Johnson[65], Rudolf Kaaks[12], Elza K. Khusnutdinova[13,94], Cari M. Kitahara[95], Stella Koutros[3], Vessela N. Kristensen[38,48], James V. Lacey[96,97], Diether Lambrechts[98,99], Flavio Lejbkowicz[100], Annika Lindblom[101,102], Michael Lush[2], Arto Mannermaa[79,103,104], Dimitrios Mavroudis[105], Usha Menon[69], Rachel A. Murphy[106,107], Heli Nevanlinna[108], Nadia Obi[109,110], Kenneth Offit[111,112], Tjoung-Won Park-Simon[20], Alpa V. Patel[18], Cheng Peng[58], Paolo Peterlongo[113], Guillermo Pita[71], Dijana Plaseska-Karanfilska[92], Katri Pylkäs[114,115], Paolo Radice[116], Muhammad U. Rashid[77,117], Gad Rennert[118], Eleanor Roberts[85], Juan Rodriguez[52], Atocha Romero[119], Efraim H. Rosenberg[120], Emmanouil Saloustros[121], Dale P. Sandler[122], Elinor J. Sawyer[123], Rita K. Schmutzler[74,75,124], Christopher G. Scott[17], Xiao-Ou Shu[125], Melissa C. Southey[126,127,128], Jennifer Stone[84,129], Jack A. Taylor[122,130], Lauren R. Teras[18], Irma van de Beek[131], Walter Willett[24,58,59], Robert Winqvist[114,115], Wei Zheng[125], Celine M. Vachon[132], Marjanka K. Schmidt[133,134,135], Per Hall[52,136], Robert J. MacInnis[84,128], Roger L. Milne[84,126,128], Paul D. P. Pharoah[137], Jacques Simard[138], Antonis C. Antoniou[2], Douglas F. Easton[2,56] and Kyriaki Michailidou[1,2*]*Correspondence:

Kyriaki Michailidou
kyriakimi@cing.ac.cy
[1] Biostatistics Unit, The Cyprus Institute of Neurology and Genetics, 6 Iroon Avenue, 2371 Ayios Dometios, Nicosia, Cyprus
[2] Department of Public Health and Primary Care, Centre for Cancer Genetic Epidemiology, University of Cambridge, Cambridge, UK
[3] Division of Cancer Epidemiology and Genetics, Department of Health and Human Services, National Cancer Institute, National Institutes of Health, Bethesda, MD, USA
[4] Fred A, Litwin Center for Cancer Genetics, Lunenfeld-Tanenbaum Research Institute of Mount Sinai Hospital, Toronto, ON, Canada
[5] Department of Molecular Genetics, University of Toronto, Toronto, ON, Canada
[6] Department of Medicine, Genetic Epidemiology Research Institute, University of California Irvine, Irvine, CA, USA
[7] NN Alexandrov Research Institute of Oncology and Medical Radiology, Minsk, Belarus
[8] Division of Clinical Epidemiology and Aging Research, German Cancer Research Center (DKFZ), Heidelberg, Germany
[9] Department of Public Health Sciences, and Cancer Research Institute, Queen's University, Kingston, ON, Canada
[10] Oncology, Clinical Sciences in Lund, Lund University, Lund, Sweden
[11] Department of Oncology, Leuven Multidisciplinary Breast Center, Leuven Cancer Institute, University Hospitals Leuven, Leuven, Belgium
[12] Division of Cancer Epidemiology, German Cancer Research Center (DKFZ), Heidelberg, Germany
[13] Institute of Biochemistry and Genetics of the Ufa Federal Research Centre of the Russian Academy of Sciences, Ufa, Russia
[14] Petersburg State University, St. Petersburg, Russia
[15] Division of Genetics and Epidemiology, The Institute of Cancer Research, London, UK
[16] Department of Genetics and Pathology, Pomeranian Medical University, Szczecin, Poland
[17] Department of Quantitative Health Sciences, Mayo Clinic, Rochester, MN, USA
[18] Department of Population Science, American Cancer Society, Atlanta, GA, USA
[19] Department of Radiation Oncology, Hannover Medical School, Hannover, Germany
[20] Gynaecology Research Unit, Hannover Medical School, Hannover, Germany
[21] Copenhagen General Population Study, Herlev and Gentofte Hospital, Copenhagen University Hospital, Herlev, Denmark
[22] Department of Clinical Biochemistry, Herlev and Gentofte Hospital, Copenhagen University Hospital, Herlev, Denmark
[23] Faculty of Health and Medical Sciences, University of Copenhagen, Copenhagen, Denmark
[24] Department of Epidemiology, Harvard TH Chan School of Public Health, Boston, MA, USA
[25] Dr. Margarete Fischer-Bosch-Institute of Clinical Pharmacology, Stuttgart, Germany
[26] iFIT-Cluster of Excellence, University of Tübingen, Tübingen, Germany
[27] German Cancer Consortium (DKTK) and German Cancer Research Center (DKFZ), Partner Site Tübingen, Tübingen, Germany
[28] German Cancer Consortium (DKTK), German Cancer Research Center (DKFZ), Heidelberg, Germany
[29] Department of Internal Medicine and Huntsman Cancer Institute, University of Utah, Salt Lake City, UT, USA
[30] Genomic Epidemiology Group, German Cancer Research Center (DKFZ), Heidelberg, Germany
[31] Oncology and Genetics Unit, Instituto de Investigacion Sanitaria Galicia Sur-Vigo-Spain, Vigo, Spain
[32] Department of Pathology, Intermountain Healthcare, Salt Lake City, UT, USA
[33] Intermountain Biorepository, Intermountain Healthcare, Salt Lake City, UT, USA
[34] Cancer Epidemiology Group, University Cancer Center Hamburg (UCCH), University Medical Center Hamburg-Eppendorf, Hamburg, Germany
[35] Cancer Research Program, QIMR Berghofer Medical Research Institute, Brisbane, QLD, Australia
[36] Department of Pediatrics, Boston Children's Hospital and Harvard Medical School, Boston, MA, USA
[37] Department of Cancer Genetics, Institute for Cancer Research, Oslo University Hospital-Radiumhospitalet, Oslo, Norway
[38] Institute of Clinical Medicine, Faculty of Medicine, University of Oslo, Oslo, Norway
[39] Department of Research, Vestre Viken Hospital, Drammen, Norway
[40] Section for Breast- and Endocrine Surgery, Division of Surgery, Department of Cancer, Cancer and Transplantation Medicine, Oslo University Hospital-Ullevål, Oslo, Norway
[41] Department of Radiology and Nuclear Medicine, Oslo University Hospital, Oslo, Norway
[42] Department of Pathology, Akershus University Hospital, Lørenskog, Norway
[43] Department of Tumor Biology, Institute for Cancer Research, Oslo University Hospital, Oslo, Norway
[44] Department of Oncology, Division of Surgery, Cancer and Transplantation Medicine, Oslo University Hospital-Radiumhospitalet, Oslo, Norway
[45] National Advisory Unit on Late Effects After Cancer Treatment, Oslo University Hospital, Oslo, Norway
[46] Department of Oncology, Akershus University Hospital, Lørenskog, Norway
[47] Oslo Breast Cancer Research Consortium, Oslo University Hospital, Oslo, Norway
[48] Department of Medical Genetics, Oslo University Hospital and University of Oslo, Oslo, Norway
[49] Department of Laboratory Medicine and Pathology, Mayo Clinic, Rochester, MN, USA
[50] Division of Clinical Medicine, School of Medicine and Population Health, University of Sheffield, Sheffield, UK
[51] Division of Neuroscience, School of Medicine and Population Health, University of Sheffield, Sheffield, UK
[52] Department of Medical Epidemiology and Biostatistics, Karolinska Institutet, Stockholm, Sweden
[53] Department of Clinical Genetics, Fox Chase Cancer Center, Philadelphia, PA, USA
[54] Department of Pathology, Leiden University Medical Center, Leiden, The Netherlands
[55] Department of Human Genetics, Leiden University Medical Center, Leiden, The Netherlands
[56] Department of Oncology, Centre for Cancer Genetic Epidemiology, University of Cambridge, Cambridge, UK
[57] Faculty of Medicine, University of Southampton, Southampton, UK
[58] Channing Division of Network Medicine, Department of Medicine, Brigham and Women's Hospital and Harvard Medical School, Boston, MA, USA
[59] Department of Nutrition, Harvard TH Chan School of Public Health, Boston, MA, USA
[60] Institute for Medical Informatics, Statistics and Epidemiology, University of Leipzig, Leipzig, Germany
[61] LIFE - Leipzig Research Centre for Civilization Diseases, University of Leipzig, Leipzig, Germany
[62] Division of Evolution and Genomic Sciences, School of Biological Sciences, Faculty of Biology, Medicine and Health, University of Manchester, Manchester Academic Health Science Centre, Manchester, UK
[63] North West Genomics Laboratory Hub, Manchester Centre for Genomic Medicine, St Mary's Hospital, Manchester University NHS Foundation Trust, Manchester Academic Health Science Centre, Manchester, UK
[64] Department of Gynecology and Obstetrics, Comprehensive Cancer Center Erlangen-EMN, Friedrich-Alexander University Erlangen-Nuremberg, University Hospital Erlangen, Erlangen, Germany
[65] The Breast Cancer Now Toby Robins Research Centre, The Institute of Cancer Research, London, UK
[66] Department of Breast Surgery, Herlev and Gentofte Hospital, Copenhagen University Hospital, Herlev, Denmark
[67] School of Population Health, Curtin University, Perth, WA, Australia
[68] Cancer Genetics and Epidemiology Group, Genomic Medicine Group, Fundación Instituto de Investigación Sanitaria de Santiago de Compostela (FIDIS), Complejo Hospitalario Universitario de Santiago, SERGAS, Santiago de Compostela, Spain

[69] MRC Clinical Trials Unit, Institute of Clinical Trials and Methodology, University College London, London, UK
[70] Department of Women's Cancer, Elizabeth Garrett Anderson Institute for Women's Health, University College London, London, UK
[71] Human Genotyping Unit-CeGen, Spanish National Cancer Research Centre (CNIO), Madrid, Spain
[72] Spanish Network on Rare Diseases (CIBERER), Madrid, Spain
[73] Team 'Exposome and Heredity', CESP, Gustave Roussy, INSERM, University Paris-Saclay, UVSQ, Villejuif, France
[74] Center for Familial Breast and Ovarian Cancer, Faculty of Medicine and University Hospital Cologne, University of Cologne, Cologne, Germany
[75] Center for Integrated Oncology (CIO), Faculty of Medicine and University Hospital Cologne, University of Cologne, Cologne, Germany
[76] Department of Preventive Medicine, Keck School of Medicine, University of Southern California, Los Angeles, CA, USA
[77] Molecular Genetics of Breast Cancer, German Cancer Research Center (DKFZ), Heidelberg, Germany
[78] Cancer RC, University of Eastern Finland, Kuopio, Finland
[79] Institute of Clinical Medicine, Pathology and Forensic Medicine, University of Eastern Finland, Kuopio, Finland
[80] Health Innovation and Evaluation Hub, Université de Montréal Hospital Research Centre (CRCHUM), Montréal, QC, Canada
[81] Department of Medical Oncology, Erasmus MC Cancer Institute, 3015 GD Rotterdam, The Netherlands
[82] Department of Gynecology and Obstetrics, University Hospital Düsseldorf, Heinrich-Heine University Düsseldorf, Düsseldorf, Germany
[83] University of Tübingen, Tübingen, Germany
[84] Centre for Epidemiology and Biostatistics, Melbourne School of Population and Global Health, The University of Melbourne, Melbourne, VIC, Australia
[85] Division of Cancer Sciences, Faculty of Biology, Medicine and Health, University of Manchester, Manchester Academic Health Science Centre, Manchester, UK
[86] Nightingale/Prevent Breast Cancer Centre, Wythenshawe Hospital, Manchester University NHS Foundation Trust, Manchester, UK
[87] Manchester Breast Centre, Manchester Cancer Research Centre, The Christie Hospital, Manchester, UK
[88] Division of Cancer Sciences, University of Manchester, Manchester, UK
[89] Australian Breast Cancer Tissue Bank, Westmead Institute for Medical Research, University of Sydney, Sydney, NSW, Australia
[90] Research Department, Peter MacCallum Cancer Center, Melbourne, VIC, Australia
[91] Sir Peter MacCallum Department of Oncology, The University of Melbourne, Parkville, VIC, Australia
[92] Research Centre for Genetic Engineering and Biotechnology 'Georgi D. Efremov', MASA, Skopje, Republic of North Macedonia
[93] Independent Laboratory of Molecular Biology and Genetic Diagnostics, Pomeranian Medical University, Szczecin, Poland
[94] Department of Genetics and Fundamental Medicine, Ufa University of Science and Technology, Ufa, Russia
[95] Division of Cancer Epidemiology and Genetics, National Cancer Institute, NIH, Bethesda, MD, USA
[96] Department of Computational and Quantitative Medicine, City of Hope, Duarte, CA, USA
[97] City of Hope Comprehensive Cancer Center, City of Hope, Duarte, CA, USA
[98] Laboratory for Translational Genetics, Department of Human Genetics, KU Leuven, Leuven, Belgium
[99] VIB Center for Cancer Biology, VIB, Leuven, Belgium
[100] Carmel Medical Center, Haifa, Israel
[101] Department of Molecular Medicine and Surgery, Karolinska Institutet, Stockholm, Sweden
[102] Department of Clinical Genetics, Karolinska University Hospital, Stockholm, Sweden
[103] Translational Cancer Research Area, University of Eastern Finland, Kuopio, Finland
[104] Biobank of Eastern Finland, Kuopio University Hospital, Kuopio, Finland
[105] Department of Medical Oncology, University Hospital of Heraklion, Heraklion, Greece
[106] School of Population and Public Health, University of British Columbia, Vancouver, BC, Canada
[107] Cancer Control Research, BC Cancer Agency, Vancouver, BC, Canada
[108] Department of Obstetrics and Gynecology, Helsinki University Hospital, University of Helsinki, Helsinki, Finland
[109] Institute for Occupational and Maritime Medicine, University Medical Center Hamburg-Eppendorf, Hamburg, Germany
[110] Institute for Medical Biometry and Epidemiology, University Medical Center Hamburg-Eppendorf, Hamburg, Germany
[111] Clinical Genetics Research Lab, Department of Cancer Biology and Genetics, Memorial Sloan Kettering Cancer Center, New York, NY, USA
[112] Clinical Genetics Service, Department of Medicine, Memorial Sloan Kettering Cancer Center, New York, NY, USA
[113] Genome Diagnostics Program, IFOM ETS - The AIRC Institute of Molecular Oncology, Milan, Italy
[114] Laboratory of Cancer Genetics and Tumor Biology, Translational Medicine Research Unit, Biocenter Oulu, University of Oulu, Oulu, Finland
[115] Laboratory of Cancer Genetics and Tumor Biology, Northern Finland Laboratory Centre Oulu, Oulu, Finland
[116] Unit of Predictive Medicine, Molecular Bases of Genetic Risk, Department of Experimental Oncology, Fondazione IRCCS Istituto Nazionale Dei Tumori (INT), Milan, Italy
[117] Department of Basic Sciences, Shaukat Khanum Memorial Cancer Hospital and Research Centre (SKMCH & RC), Lahore, Pakistan
[118] Technion, Faculty of Medicine and Association for Promotion of Research in Precision Medicine, Haifa, Israel
[119] Medical Oncology Department, Hospital Universitario Puerta de Hierro, Madrid, Spain
[120] Department of Pathology, The Netherlands Cancer Institute - Antoni Van Leeuwenhoek Hospital, Amsterdam, The Netherlands
[121] Department of Oncology, University Hospital of Larissa, Larissa, Greece
[122] Epidemiology Branch, National Institute of Environmental Health Sciences, NIH, Research Triangle Park, NC, USA
[123] School of Cancer and Pharmaceutical Sciences, Comprehensive Cancer Centre, Guy's Campus, King's College London, London, UK
[124] Center for Molecular Medicine Cologne (CMMC), Faculty of Medicine and University Hospital Cologne, University of Cologne, Cologne, Germany
[125] Division of Epidemiology, Department of Medicine, Vanderbilt Epidemiology Center, Vanderbilt-Ingram Cancer Center, Vanderbilt University School of Medicine, Nashville, TN, USA
[126] Precision Medicine, School of Clinical Sciences at Monash Health, Monash University, Clayton, VIC, Australia
[127] Department of Clinical Pathology, The University of Melbourne, Melbourne, VIC, Australia
[128] Cancer Epidemiology Division, Cancer Council Victoria, Melbourne, VIC, Australia
[129] Genetic Epidemiology Group, School of Population and Global Health, University of Western Australia, Perth, WA, Australia
[130] Epigenetic and Stem Cell Biology Laboratory, National Institute of Environmental Health Sciences, NIH, Research Triangle Park, NC, USA
[131] Department of Clinical Genetics, The Netherlands Cancer Institute - Antoni Van Leeuwenhoek Hospital, Amsterdam, The Netherlands
[132] Division of Epidemiology, Department of Quantitative Health Sciences, Mayo Clinic, Rochester, MN, USA
[133] Division of Molecular Pathology, The Netherlands Cancer Institute, Amsterdam, The Netherlands
[134] Division of Psychosocial Research and Epidemiology, The Netherlands Cancer Institute - Antoni Van Leeuwenhoek Hospital, Amsterdam, The Netherlands
[135] Department of Clinical Genetics, Leiden University Medical Center, Leiden, The Netherlands
[136] Department of Oncology, Södersjukhuset, Stockholm, Sweden
[137] Department of Computational Biomedicine, Cedars-Sinai Medical Center, West Hollywood, CA, USA
[138] Genomics Center, Centre Hospitalier Universitaire de Québec – Université Laval Research Center, Québec City, QC, Canada

Full list of author information is available at the end of the article

Abstract

Background The 313-variant polygenic risk score (PRS_{313}) provides a promising tool for clinical breast cancer risk prediction. However, evaluation of the PRS_{313} across different European populations which could influence risk estimation has not been performed.

Methods We explored the distribution of PRS_{313} across European populations using genotype data from 94,072 females without breast cancer diagnosis, of European-ancestry from 21 countries participating in the Breast Cancer Association Consortium (BCAC) and 223,316 females without breast cancer diagnosis from the UK Biobank. The mean PRS was calculated by country in the BCAC dataset and by country of birth in the UK Biobank. We explored different approaches to reduce the observed heterogeneity in the mean PRS across the countries, and investigated the implications of the distribution variability in risk prediction.

Results The mean PRS_{313} differed markedly across European countries, being highest in individuals from Greece and Italy and lowest in individuals from Ireland. Using the overall European PRS_{313} distribution to define risk categories, leads to overestimation and underestimation of risk in some individuals from these countries. Adjustment for principal components explained most of the observed heterogeneity in the mean PRS. The mean estimates derived when using an empirical Bayes approach were similar to the predicted means after principal component adjustment.

Conclusions Our results demonstrate that PRS distribution differs even within European ancestry populations leading to underestimation or overestimation of risk in specific European countries, which could potentially influence clinical management of some individuals if is not appropriately accounted for. Population-specific PRS distributions may be used in breast cancer risk estimation to ensure predicted risks are correctly calibrated across risk categories.

Keywords Polygenic risk scores, Breast cancer, Risk prediction, Risk calibration

Background

Genetic susceptibility to breast cancer is influenced by multiple genetic variants that contribute to different levels of risk [1–6]. Genome-wide association studies (GWAS) have identified a large number of common variants that each contribute a small risk to the disease but can be combined into polygenic risk scores (PRSs) with greater effects [7, 8]. PRSs provide a promising tool for clinical breast cancer risk prediction by stratifying women into different risk categories [9–11] and may be used to inform targeted screening and prevention strategies [12–20].

Mavaddat et al. [11] constructed a 313-variant PRS (PRS_{313}) for breast cancer using data from women of European ancestry participating in the Breast Cancer Association Consortium (BCAC). In prospective validation studies, this PRS was estimated to be associated with a relative risk for breast cancer of ~ 1.6 per standard deviation (SD) increase. The lifetime absolute risk of developing overall breast cancer for women in the 1% of the PRS_{313} risk distribution was ~ 2%; while for those in the 99% was 32.6%. PRS_{313} has been incorporated into the CanRisk tool (www.canrisk.org) [14, 21, 22] and together with other lifestyle and genetic risk factors, has been shown to improve risk stratification in European ancestry populations [14, 23–27]. Several large studies have investigated the transferability of PRSs developed in European ancestry population to non-European populations, finding that the strength of associations with breast cancer risk were attenuated, particularly among women of African ancestry, compared to association among women of European ancestry [28–30].

PRS distributions across different European countries have not, however, been extensively evaluated. Differences in the PRS distribution, if not appropriately accounted for, could lead to inappropriate risk classification, with implications for clinical management. Here, we examined the distribution of the PRS_{313} across 17 countries in Europe, together with individuals of European ancestry from Australia, Canada, Israel and the USA. Similar analyses were performed using data from the UK Biobank, stratifying individuals by country of birth. We explored different approaches to account for PRS_{313} distribution differences across countries, and investigated the implications of the observed variability for breast cancer risk prediction.

Methods

Study populations

Breast Cancer Association Consortium dataset

The BCAC dataset used here consisted of 110,260 female invasive breast cancer cases and 94,072 female controls of European ancestry who were recruited into 84 studies from 21 countries participating in the BCAC (Table S1A). For simplicity and in an attempt to explore the effect on

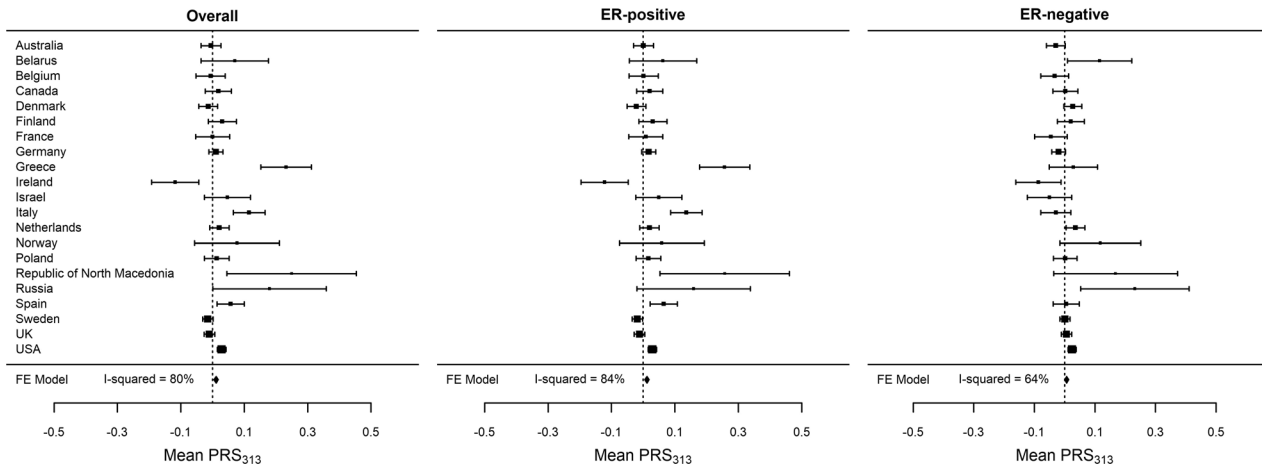

Fig. 1 Standardized PRS_{313} distribution across countries for overall, ER-positive and ER-negative breast cancer in BCAC. The squares represent the mean PRS by country, and the error bars represent the corresponding 95% confidence intervals. ER, Oestrogen receptor; FE Model, Fixed-effects Model; PRS, Polygenic risk score

Table 1 Mean standardized PRS_{313} by country in controls of the pooled BCAC dataset

Country	Number of controls	Mean PRS_{313}[a]	Mean PRS adjusted for array and 6 PCs[b]	PRS adjusted for 6 PCs, fitted values[c]	Empirical Bayes Posterior Mean[d]
Australia	4049	−0.005	0.01	−0.005	−0.003
Belarus	342	0.07	0.071	0.016	0.064
Belgium	1823	−0.006	−0.007	0.010	−0.002
Canada	2277	0.018	0.019	0.013	0.02
Denmark	5241	−0.013	0.012	−0.031	−0.012
Finland	2083	0.031	0.008	0.010	0.032
France	1372	0.0003	−0.008	0.008	0.004
Germany	8563	0.011	0.004	0.013	0.011
Greece	607	0.232	0.043	0.208	0.199
Ireland	719	−0.118	−0.015	−0.112	−0.092
Israel	724	0.047	0.001	0.062	0.047
Italy	1554	0.115	−0.007	0.131	0.11
Netherlands	4407	0.021	0.043	−0.019	0.022
Norway	217	0.077	0.094	−0.027	0.066
Poland	2554	0.013	0.025	0.010	0.015
Republic of North Macedonia	92	0.25	0.134	0.140	0.129
Russia	120	0.18	0.166	0.044	0.11
Spain	2098	0.057	−0.006	0.057	0.056
Sweden	16,680	−0.015	0.005	−0.017	−0.014
UK	16,854	−0.01	0.019	−0.023	−0.01
USA	21,696	0.029	0.033	0.013	0.029

This table presents the mean PRS by country when adjusted for array type, 6 PCs and array type, using fitted values adjusted for 6 PCs and array type, and when using an empirical Bayes approach adjusted for array type

[a] Mean PRS_{313} after adjustment for array type

[b] Mean PRS after adjustment for 6 PCs. Since PC adjustment explains much of the variation, these values are typically close to 0

[c] Mean PRS_{313} by country based on averaging the predicted PRS of each individual; estimated using linear predictor of PRS versus 6 PCs and the command predict () in R

[d] Country-specific estimates, means β, using the empirical Bayes approach, adjusted for array type

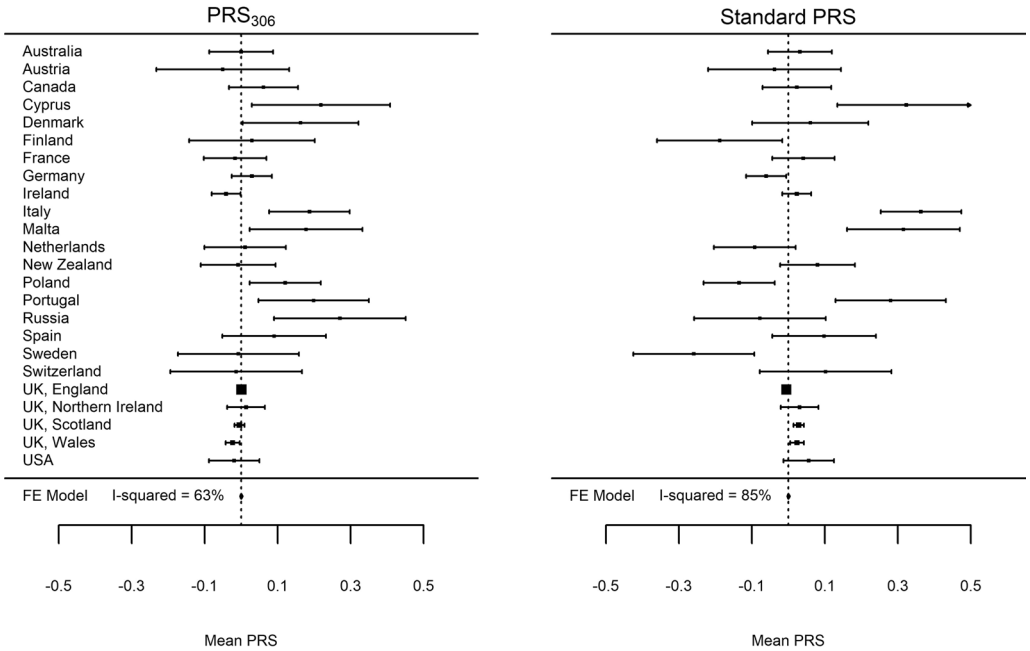

Fig. 2 PRS distribution across countries for overall breast cancer in the UK Biobank. Distribution of the mean PRS_{306} and "standard" PRS for breast cancer, as defined in the UK Biobank, across countries of origin for participating white females. The squares represent the mean PRS by country, and the error bars represent the corresponding 95% confidence intervals. FE Model, Fixed-effects Model; PRS, Polygenic risk score

the general female population, only the control data were used. Samples were genotyped using the iCOGS [1] or OncoArray [3, 31] genotyping arrays. The iCOGS and OncoArray datasets were imputed separately and ancestry-informative principal components (PCs) were calculated, as described previously [2, 3, 31].

UK Biobank dataset

Genotype data from females (genetically reported sex) participating in the UK Biobank were used. Individuals were excluded if they had a recorded breast cancer diagnosis (malignant neoplasm or carcinoma in situ of the breast) or had a personal history of malignant neoplasm of the breast, based on the cancer registry or self-reported. Individuals with a SNP call rate < 0.95 were removed from the analysis. Genetic ancestry was inferred using FastPop software [32]. Individuals self-reported as "white" and with an estimated European ancestry proportion ≥ 80% were retained in the analysis. Individuals were subsequently stratified by the "country of birth" field in the UK Biobank; only countries with at least 100 participants were included. After filtering, 223,316 females from 21 countries were included in the analyses (Table S1B). More details on the genotyping, quality control, imputation procedures used, and calculation of PCs are given elsewhere [33, 34].

All participants provided written informed consent, and all the studies were approved by the relevant ethics committees. The use of UK Biobank data has been approved under the application with ID102655, and BCAC data under the application with access number 712.

Statistical analysis

PRS_{313} was developed previously [11] and included variants independently associated with breast cancer risk at a P cut-off $< 10^{-05}$. The PRS_{313} was calculated for each study participant using the following formula:

$$PRS_j = \beta_1 x_{j1} + \cdots \beta_k x_{jk} + \beta_{313} x_{j;313}$$

where PRS_j is the PRS of individual j, x_{jk} is the estimated effect allele dosage for SNP_k carried by individual j and can take values between 0 and 2, and β_k is the weight for SNP_k in the PRS for overall breast cancer, as derived by Mavaddat et al. [11] PRS_{313} was standardized to have unit SD in controls in the pooled dataset. Mavaddat et al. also derived specific versions of PRS_{313} for oestrogen receptor (ER) subtypes, with weights optimised for predicting ER-positive or ER-negative breast cancer risk (Table S2). The main analyses focused on calculating the mean standardized PRS_{313} in BCAC controls using both the iCOGS and OncoArray datasets. These values were derived using

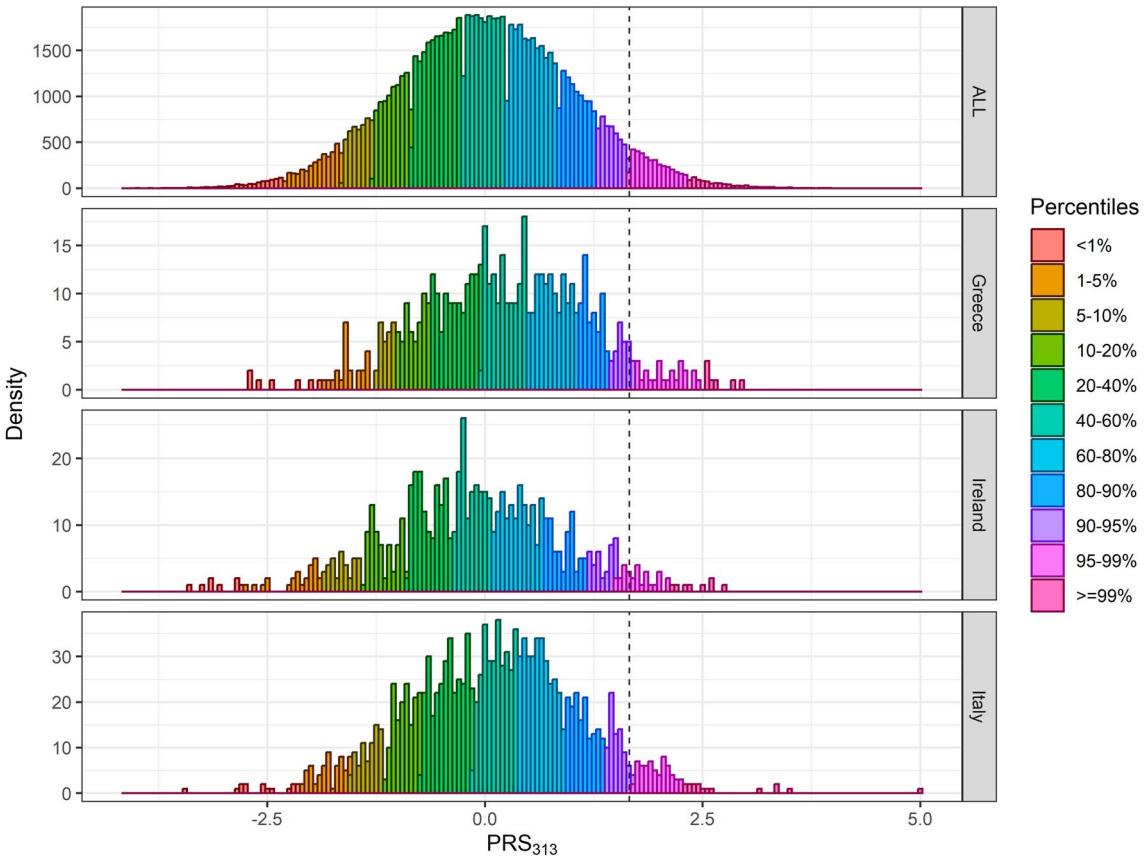

Fig. 3 PRS_{313} distribution by percentiles in the pooled BCAC dataset, Greece, Ireland and Italy. The dashed line corresponds to the 95th percentile of the PRS_{313} distribution in controls of the pooled BCAC dataset

Table 2 Risk estimation examples using the CanRisk tool

Samples used for the standardization	Mean (SD)	Standardized PRS[a]	Percentage based on CanRisk tool (%)	Lifetime risk based on CanRisk tool[b] (%)	NICE Risk category[c]
Individual from Greece with raw PRS_{313} = 0.34 (falling into the 90–95% percentile category in the full BCAC dataset)					
CanRisk tool[d]	−0.424 (0.611)	1.253	89.5	14.1	Moderate
Controls Greece (raw)[e]	−0.305 (0.612)	1.056	85.5	13.3	Population
Controls Greece adjusted for 6 PCs (raw)[e]	−0.420 (0.696)	1.094	86.3	13.5	Population
Controls Greece, using Empirical Bayes method[e]	−0.325 (0.554)	1.204	88.6	13.9	Moderate
Individual from Ireland with raw PRS_{313} = 0.27 (falling into the 85–90% percentile category in the full BCAC dataset)					
CanRisk tool[d]	−0.424 (0.611)	1.14	87.3	13.7	Population
Controls Ireland (raw)[e]	−0.519 (0.624)	1.27	89.8	14.2	Moderate
Controls Ireland adjusted for 6 PCs (raw)[e]	−0.456 (0.74)	0.985	83.8	13	Population
Controls Ireland, Empirical Bayes method[e]	−0.503 (0.562)	1.38	91.7	14.7	Moderate

Mean and SD used to standardize PRS_{313} of a 50-year-old woman with a raw PRS_{313} equal to 0.34 from Greece and another 50-year-old woman with a raw PRS_{313} equal to 0.27 from Ireland; the risk estimation and categorization were performed when using the CanRisk tool

[a] Standardized based on the mean and SD specified in the second column

[b] Absolute risk of developing breast cancer by the age of 80

[c] These breast cancer risk categories provided from the CanRisk tool are based on the risk assessment and according to the National Institute for Health and Care Excellence guideline (NICE-CG164) on familial breast cancer and correspond to "Near population risk", "Moderate risk" and "High risk"

[d] When a variant call format (vcf) file is uploaded to the CanRisk tool, a raw PRS_{313} can be calculated and standardized using the mean (SD): −0.424 (0.611)

[e] Adjusted for array type

linear regression with array type as a covariate and no intercept (so that estimates were generated for every country). Heterogeneity in the mean PRS_{313} between countries was assessed using I^2 statistics and Q statistic *P*-values.

We also evaluated the distribution of the mean PRS by country of birth in the UK Biobank dataset. Seven of the 313 variants were not available in the UK Biobank data; thus, we used the remaining 306 variants in the analysis (PRS_{306}) (Table S2). PRS_{306} was standardized to have unit SD in controls in the pooled UK Biobank dataset. We also evaluated a "standard" breast cancer PRS available in the UK Biobank data, previously generated from external GWAS data [35] and was available for 222,989 individuals (Table S1B). This PRS was also standardized to have unit SD in controls in the pooled UK Biobank dataset.

Potential sources of the variability in the mean PRS_{313} across the countries were explored in the BCAC dataset using three approaches. The PRS was first recalculated excluding variants in the *CHEK2* region. The protein truncating variant *CHEK2* c.1100delC is a relatively common founder variant that exhibits a large variation in frequency across Europe [36]. Although it is not included in PRS_{313}, other variants in PRS_{313} are correlated with this variant (Table S2) and were removed.

Second, we examined the effect of removing variants with the most variable frequency across countries. The mean and SD of the effect allele frequency across countries, in controls of the pooled dataset were calculated for each of the 313 variants. Variants with a coefficient of variation (SD/mean) > 0.3 were removed.

Third, we explored the effect of adjusting for up to 10 ancestry-informative PCs, in addition to array type. As the PCs derived from the iCOGS and OncoArray databases are not comparable, separate PCs for each were included in the regression. We explored the number of PCs that were required to eliminate heterogeneity in the adjusted mean PRS_{313} using the thresholds $I^2 < 10\%$ and $P > 0.05$. Similarly, for the UK Biobank dataset, PRS_{306} was adjusted for up to 10 PCs, which were available in the UK Biobank.

As a complementary approach to generating population-specific estimates, we explored an empirical Bayes approach similar to that described by Clayton and Kaldor [37] for mapping disease rates (details in Additional File 1).

To investigate the implications of PRS distribution differences in breast cancer risk prediction, we explored the proportion of women by country by percentile based on the distribution cut-offs of either the full dataset or country-specific values, separately in the BCAC and the UK Biobank. We also examined two specific risk estimation examples using the CanRisk tool [14, 21, 22].

All analyses were performed in R (version 4.2.1).

Results

Geographic diversity in the mean PRS_{313} across European ancestry populations

The mean PRS_{313} in the BCAC controls differed markedly across European countries, with $I^2 = 80\%$ ($P = 5.6 \times 10^{-13}$). The mean was highest in the Republic of North Macedonia and Greece and lowest in Ireland. A similar level of heterogeneity was observed for the ER-positive ($I^2 = 84\%$) and ER-negative ($I^2 = 64\%$) PRSs. There was no evidence of a difference in the SD of the PRS between countries (Fig. 1; Tables 1, S3).

The mean PRS_{306} in female UK Biobank participants, stratified by country of birth, was also calculated. There was strong evidence of heterogeneity in the PRS distribution ($I^2 = 63\%$, $P = 1.7 \times 10^{-05}$). The pattern was generally similar to that seen in the BCAC dataset, with a higher mean PRS in individuals born in Cyprus, Russia, and Italy) and a lower PRS in Ireland). Similar results were found for the "standard" UK Biobank PRS ($I^2 = 85\%$, $P = 8.5 \times 10^{-21}$) (Fig. 2; Table S4).

Exploring potential reasons for differences in the mean PRS between countries

Potential sources of the variability in the mean PRS_{313} across the countries were explored in the BCAC dataset using three approaches. After removing variants in the *CHEK2* region, the variation in the mean PRS across countries remained similar to PRS_{313} ($I^2 = 83\%$, $P = 9.4 \times 10^{-16}$). We next identified the variants with the most variable frequency among the countries. Seventeen variants had a coefficient of variation > 0.3 (Table S2). Excluding these 17 variants did not reduce the variation in the mean PRS ($I^2 = 80\%$, $P = 2.4 \times 10^{-12}$).

We next explored the effect of adjusting for PCs. When individuals in the BCAC dataset genotyped with OncoArray were plotted by the first two PCs, those from the same country separated clearly in a pattern consistent with their geographical relationship (Fig. S1). This finding suggested that adjusting for PCs maybe an effective approach for reducing the variation in PRS distribution. When we adjusted the PRS for the leading PCs in the BCAC dataset, the I^2 decreased as each PC was added to the model and reached < 10% when adjusted for the first six PCs (Table 1, Table S3, Fig. S2). A similar result was obtained for the ER-positive PRS, after adjustment for the first six PCs ($I^2 = 0\%$, $P = 0.69$). For the ER-negative PRS, however, heterogeneity was not eliminated even when the PRS was adjusted for 10 PCs ($I^2 = 56\%$, $P = 0.001$) (Table S3). The predicted PRS of each individual, as derived from the fitted values of the linear regression model of PRS adjusted for the first six

PCs and array type, was subsequently used to calculate a predicted mean PRS_{313} by country (Tables 1, S3). We repeated these analyses for PRS_{306} using the UK Biobank dataset. I^2 decreased as each PC was added to the model and reached < 10% and ~ 0% when adjusted for the first seven and eight PCs, respectively (Fig. S3, Table S4).

Mean PRS estimates by country calculated using an empirical Bayes approach

The empirical Bayes estimates by country for the mean PRS were calculated in the BCAC dataset (Table 1, Table S5). Compared with the unadjusted estimates, the estimates shrunk toward the overall mean, with shrinkage being greatest for countries with small available sample sizes. The adjusted mean PRS by country were generally similar to those predicted by the model adjusted for six PCs. When PRSs were adjusted for the first six PCs, applying the empirical Bayes approach made little difference in the estimates.

Implications for Breast Cancer Risk Prediction

To explore the effect of PRS distribution differences among European populations on risk stratification, we first defined risk thresholds based on the distribution of the controls in the full BCAC and the UK Biobank datasets separately. We then calculated the percentage of controls by country that would be categorized in each percentile based on the distribution in the full dataset and compared these to the percentages based on the country-specific distributions (Tables S6, S7, S8). PRS_{313} percentile distribution in the full BCAC dataset, Greece, Italy (highest PRS_{313} and including > 100 controls) and Ireland (lowest PRS_{313}) are illustrated (Fig. 3, Table S7). Based on the overall distribution, ~ 1.3% and ~ 0.5% additional women from Greece, and Italy, respectively, were incorrectly classified in the 95–99th percentile instead of in the 90–95th percentile, while ~ 1.4% additional women from Ireland were incorrectly classified in the 90–95th instead of the 95–99th percentile (Table S6C). Similar results were observed for the UK Biobank (Fig. S4).

An example a 50-year-old female from Greece with a raw PRS_{313} of 0.34 (falling into the 90–95th percentile-in the full BCAC dataset) and no other risk factors known was considered. Using the CanRisk tool she would be classified in the moderate risk category. If the PRS were standardized based on the mean and SD of the controls from Greece or based on the values of PRS for Greece predicted by adjustment for the first six PCs, she would be classified into the population risk category. If the PRS were standardized based on the values of the empirical Bayes approach she would be classified into the moderate risk category (Table 2).

A second example based on a 50-year-old female from Ireland with a raw PRS_{313} equal to 0.27 (falling into the 85–90th percentile-in the full BCAC dataset), and no other risk factors known was considered. Using the CanRisk tool, she would be classified in the population risk category. If the PRS were standardized based on the mean and SD of PRS_{313} as derived from the controls in Ireland or based on the values of the empirical Bayes approach, she would be classified in the moderate risk category. If the PRS was standardized based on values of PRS for Ireland predicted by adjustment for the first six PCs, she would be classified in the population risk category (Table 2).

Discussion

The transferability of PRSs across different populations remains a major challenge in the field of personalized cancer risk prediction [38, 39]. Here, we explored the distribution of PRS_{313} for breast cancer in women of European ancestry from 21 countries using data from studies participating in the BCAC and further investigated how the observed variability might be accounted for in breast cancer risk prediction.

The results indicated that the PRS_{313} distribution varies markedly even within European ancestry populations, with a higher mean in Greece and Italy and a lower mean in Ireland. We observed a very similar pattern in females participating in the UK Biobank based on country of birth. If not accounted for, these differences could lead to an over- or underestimation of risk, thus affecting the risk categorization and possibly the clinical management of some women. This may be important not only at the individual country level but also for individuals living in a different country than their origin.

The variability in the mean PRS_{313} could not be explained by removing variants with the most variable frequency, indicating that a large number of variants may contribute to this difference. Removing such variants to reduce heterogeneity would not be desirable, as it would reduce the risk discrimination provided by the PRS. The results do, however, indicate that most, if not all, of the variability in the mean PRS_{313} across countries in controls can be explained by adjusting for the leading ancestry-informative PCs.

We also explored generating country-specific mean PRS using an empirical Bayes approach. This approach considers both the uncertainty due to the small sample size and the true variation in the means across the countries; these country-specific mean PRSs were similar to those generated by adjusting for PCs. These values can then be used to standardize the PRS before, for example, it is implemented in the CanRisk tool. CanRisk is an online tool that enables healthcare professionals to

calculate an individual's future risk of developing breast and ovarian cancer using a combination of genetic factors (including the PRS), lifestyle/hormonal risk factors, breast density and family history. The risks are provided both over a period of time (e.g. 10 years) and lifetime, and these risks can be used to classify an individual according to management guidelines, including the National Institute for Health and Care Excellence guideline (NICE-CG164) on familial breast cancer (which classifies individuals as "Near population risk", "Moderate risk" and "High risk") [40].

The optimal approach to calibration will depend on what data are available. If a large control sample (n > 1,000) is available, it will be preferable to utilise estimates from this. If sample sizes are smaller, there seems little to choose between adjustment for PCs or an empirical Bayes approach. Adjustment for PCs has the advantage into account spatial variation. Using PCs has the advantage that they do not require any prior data from the population in question, and the approach naturally takes into account spatial variation in the PRS. A disadvantage, however, is that PCs require array genotyping data to generate, making them less attractive when implemented using sequencing panels. Moreover, the PCs generated using different genotyping arrays are not necessarily comparable. We also note that the heterogeneity of the ER-negative specific PRS was not eliminated even with the adjustment for 10 PCs. The empirical Bayes approach is simpler to implement, providing some control data are available for the population of interest.

The risk categorization of the two examples when using the CanRisk tool in the Results section, was changed depending on the mean and SD of the sample used for the standardization of the PRS. According to the NICE guideline CG164, women classified in the "Moderate risk" category have different managing guidelines than women classified in the "Near population risk" category [40].

While adjustment of the PRS distribution at the population level is clearly necessary, the results raise the question as to whether it is appropriate in general to adjust the PRS for PCs at the individual level, which gives different scores and potentially different risk classifications. This is a difficult question to address and hinges on whether the PCs should be regarded as nuisance parameters correcting for confounding factors, such as screening or lifestyle factors. Reanalysis of prospective studies with the BCAC OncoArray dataset showed that the first two PCs are associated with the PRS (PC1 negatively, PC2 positively) and are also associated with risk (in the same direction). The PRS effect size (OR per 1 SD) was essentially unchanged whether or not adjustment was made for PCs (data not shown). This finding implies that risk discrimination could be slightly improved by including the effect of PCs in the PRS and that adjusting the PRS for PCs further reduces the discrimination ability. Fortunately, the association between PC1 and risk is weak, and within a country, the variation in PC1 is not large enough to materially change risk categories.

The differences in the PRS distribution across Europe are a manifestation, on a continental scale, of the larger intercontinental differences—the mean PRS is higher in both East Asian and African populations than in the European dataset examined here [28, 29, 41]. Interestingly, the pattern within European ancestry women appears to be unrelated to population-specific incidence which is lower in Italy and Greece than in north-western Europe, including Ireland, UK, and Scandinavia [42], presumably because the effect on disease incidence is counterbalanced by greater effects of lifestyle (or other genetic) factors. It remains unclear whether the differences in PRS can be attributed purely to random genetic drift or whether selection pressures relevant to breast cancer aetiology are involved.

We should emphasise that, while adjustment for the PRS distribution is clearly important, there is no evidence for variation in the effect size (relative risk per standard deviation). Different effect sizes could result from different variant allele frequencies and (since most of the SNPs in the PRS are not causal) differences in linkage disequilibrium patterns. However, there is no evidence for this—the effect sizes (relative risks per standard deviation) are very similar across prospective validation studies [11, 26], though there is admittedly not yet good prospective data for southern/eastern-European populations. Whilst attenuation of the effect size is seen in non-European populations, the any different in effect size among European populations is likely to be very small.

We would like to acknowledge several potential limitations of our study. The dataset we used was genetically homogeneous and may not be completely representative of the population of each country. How to interpret the PRS in individuals classified as mixed ancestry is an important issue that could be explored. Furthermore, evaluation of the country-specific calibrated PRS in combination with classical breast cancer risk factors should be performed to explore the ability of these findings to predict the final risk. Finally, while we have evaluated the variation in PRS among European populations, similar issues will apply to PRS in other ancestries and in other countries, and to groups of more mixed ancestry. Similar approaches, using a combination of population-specific control data, principal component adjustment and/or empirical Bayes estimation, should also be useful for PRS calibration more generally.

In summary, these results demonstrate that the implementation of the PRS_{313} in risk prediction models such as CanRisk/BOADICEA could require country-specific calibration. This can be achieved by genotyping a large control group to obtain population-specific means, by using a PC adjustment, or the empirical Bayes approach described here.

Conclusions

In this study, we observed a remarkable difference in the mean breast cancer PRS within European ancestry populations, when we used data from more than 300,000 women with no previous breast cancer diagnosis. This heterogeneity could influence the classification of some individuals if not appropriately accounted for, leading to risk overestimation in some individuals and risk underestimation inothers, with potential implications for clinical management. Adjusting for principal components seems to correct distribution differences across populations. Therefore, the implementation of PRS for breast cancer risk prediction in European ancestry populations, will required population-specific calibration, for more accurate risk estimation. This is particularly important for countries not represented in the original PRS development.

Abbreviations

BCAC	Breast Cancer Association Consortium
BOADICEA	The Breast and Ovarian Analysis of the Disease Incidence and Carrier Estimation Algorithm
COGS	Collaborative Oncological Gene-Environment Study
ER	Oestrogen receptors
FE Model	Fixed-effects Model
GWAS	Genome-wide association studies
OR	Odds ratio
P	*P*-Value
PC	Principal component
PRS	Polygenic risk score
SD	Standard deviation
SE	Standard error

Supplementary Information

The online version contains supplementary material available at https://doi.org/10.1186/s13058-024-01947-x.

Additional file 1.

Additional file 2.

Acknowledgements

This research has been conducted using the UK Biobank Resource under application number 102655. The remaining acknowledgements are available online.

Author contributions

Writing Group: KY, KMi, DFE, ACA, NMa, and JSi; Study design: KMi, DFE, ACA, NMa, JSi, and KY; Data management: MKB, QW; Statistical Analysis: KY, NMa, JD, MZ, DFE, KMi; Provided data: MA, TUA, ILA, HA-C, NNA, VA, KJA, AAu, ABat, SBe, MBerm, ABer, KBia, NB, CBo, NVB, SEB, KBr, HBra, HBre, NJC, FC, JEC, JC-C, GC-T, WKC, NBCS Collaborators, SVC, FJC, ACox, SSC, KCz, MBD, PD, TD, AMD, DME, AHE, CEn, ME, DGE, PAF, OF, HF, MG-D, AG-M, AG-N, PGu, EHah, CAH, PHall, UH, JMH, VH, JH, AHol, EHon, MJH, RH, JLHo, SH, AHow, ABCTB Investigators, kConFab Investigators, SJ, AJak, HJ, NJ, RKa, EKK, CMKi, SKou, VNK, JVL, DLa, FLej, ALin, MLus, RJM, AMan, DM, UM, RLM, RAM, HNe, NOb, KOf, T-WP-S, AVP, CP, PPe, PDPP, GPi, DPK, KPy, PRa, MUR, GR, ER, JR, ARo, EHR, ES, DPS, EJS, MKS, RKS, CSc, X-OS, MCS, JSt, JAT, LRT, CMV, IVDB, WW, RWi, WZ, JSi, ACA, DFE. All authors read and approved the final version of the manuscript.

Funding

All funding information are available online.

Availability of data and materials

The BCAC and the UK Biobank data that support the findings of this study are available via application to the Data Access and Co-ordination Committee (BCAC@medschl.cam.ac.uk) and via application to https://www.ukbiobank.ac.uk/enable-your-research/apply-for-access, respectively. The Breast Cancer Association Consortium data have been used under the application with access number 712. The UK Biobank data have been used under the access application with ID: 102655.

Declarations

Ethics approval and consent to participate

All study participants gave written informed consent, and all the Breast Cancer Association Consortium studies were approved by the relevant ethics committees. The Breast Cancer Association Consortium data have been used under the application with access number 712. The use of the UK Biobank has been approved under application ID102655.

Consent for publication

Not applicable.

Competing interests

The following authors declare conflicts not directly relevant to this work as stated below: U.M. has a patent (no: EP10178345.4) for Breast Cancer Diagnostics and held personal shares in Abcodia Ltd between 2011 and 2021. She has research collaborations with Mercy Bioanalytics, iLOF, RNA Guardian and Micronoma in the field of early detection of cancer. P.A.F. conducts research funded by Amgen, Novartis and Pfizer. He received Honoraria from Roche, Novartis and Pfizer. R.A.M. is a Consultant for Pharmavite.

Author details

[1]Biostatistics Unit, The Cyprus Institute of Neurology and Genetics, 6 Iroon Avenue, 2371 Ayios Dometios, Nicosia, Cyprus. [2]Department of Public Health and Primary Care, Centre for Cancer Genetic Epidemiology, University of Cambridge, Cambridge, UK. [3]Division of Cancer Epidemiology and Genetics, Department of Health and Human Services, National Cancer Institute, National Institutes of Health, Bethesda, MD, USA. [4]Fred A, Litwin Center for Cancer Genetics, Lunenfeld-Tanenbaum Research Institute of Mount Sinai Hospital, Toronto, ON, Canada. [5]Department of Molecular Genetics, University of Toronto, Toronto, ON, Canada. [6]Department of Medicine, Genetic Epidemiology Research Institute, University of California Irvine, Irvine, CA, USA. [7]NN Alexandrov Research Institute of Oncology and Medical Radiology, Minsk, Belarus. [8]Division of Clinical Epidemiology and Aging Research, German Cancer Research Center (DKFZ), Heidelberg, Germany. [9]Department of Public Health Sciences, and Cancer Research Institute, Queen's University, Kingston, ON, Canada. [10]Oncology, Clinical Sciences in Lund, Lund University, Lund, Sweden. [11]Department of Oncology, Leuven Multidisciplinary Breast Center, Leuven Cancer Institute, University Hospitals Leuven, Leuven, Belgium. [12]Division of Cancer Epidemiology, German Cancer Research Center (DKFZ), Heidelberg, Germany. [13]Institute of Biochemistry and Genetics of the Ufa Federal Research Centre of the Russian Academy of Sciences, Ufa, Russia. [14]Petersburg State University, St. Petersburg, Russia. [15]Division of Genetics and Epidemiology, The Institute of Cancer Research, London, UK. [16]Department of Genetics and Pathology, Pomeranian Medical University, Szczecin, Poland. [17]Department of Quantitative Health Sciences, Mayo Clinic, Rochester, MN, USA. [18]Department of Population Science, American Cancer Society, Atlanta, GA, USA. [19]Department of Radiation Oncology, Hannover Medical School, Hannover, Germany. [20]Gynaecology Research Unit, Hannover Medical School, Hannover, Germany. [21]Copenhagen General Population Study, Herlev

and Gentofte Hospital, Copenhagen University Hospital, Herlev, Denmark. [22]Department of Clinical Biochemistry, Herlev and Gentofte Hospital, Copenhagen University Hospital, Herlev, Denmark. [23]Faculty of Health and Medical Sciences, University of Copenhagen, Copenhagen, Denmark. [24]Department of Epidemiology, Harvard TH Chan School of Public Health, Boston, MA, USA. [25]Dr. Margarete Fischer-Bosch-Institute of Clinical Pharmacology, Stuttgart, Germany. [26]iFIT-Cluster of Excellence, University of Tübingen, Tübingen, Germany. [27]German Cancer Consortium (DKTK) and German Cancer Research Center (DKFZ), Partner Site Tübingen, Tübingen, Germany. [28]German Cancer Consortium (DKTK), German Cancer Research Center (DKFZ), Heidelberg, Germany. [29]Department of Internal Medicine and Huntsman Cancer Institute, University of Utah, Salt Lake City, UT, USA. [30]Genomic Epidemiology Group, German Cancer Research Center (DKFZ), Heidelberg, Germany. [31]Oncology and Genetics Unit, Instituto de Investigacion Sanitaria Galicia Sur-Vigo-Spain, Vigo, Spain. [32]Department of Pathology, Intermountain Healthcare, Salt Lake City, UT, USA. [33]Intermountain Biorepository, Intermountain Healthcare, Salt Lake City, UT, USA. [34]Cancer Epidemiology Group, University Cancer Center Hamburg (UCCH), University Medical Center Hamburg-Eppendorf, Hamburg, Germany. [35]Cancer Research Program, QIMR Berghofer Medical Research Institute, Brisbane, QLD, Australia. [36]Department of Pediatrics, Boston Children's Hospital and Harvard Medical School, Boston, MA, USA. [37]Department of Cancer Genetics, Institute for Cancer Research, Oslo University Hospital-Radiumhospitalet, Oslo, Norway. [38]Institute of Clinical Medicine, Faculty of Medicine, University of Oslo, Oslo, Norway. [39]Department of Research, Vestre Viken Hospital, Drammen, Norway. [40]Section for Breast- and Endocrine Surgery, Division of Surgery, Department of Cancer, Cancer and Transplantation Medicine, Oslo University Hospital-Ullevål, Oslo, Norway. [41]Department of Radiology and Nuclear Medicine, Oslo University Hospital, Oslo, Norway. [42]Department of Pathology, Akershus University Hospital, Lørenskog, Norway. [43]Department of Tumor Biology, Institute for Cancer Research, Oslo University Hospital, Oslo, Norway. [44]Department of Oncology, Division of Surgery, Cancer and Transplantation Medicine, Oslo University Hospital-Radiumhospitalet, Oslo, Norway. [45]National Advisory Unit on Late Effects After Cancer Treatment, Oslo University Hospital, Oslo, Norway. [46]Department of Oncology, Akershus University Hospital, Lørenskog, Norway. [47]Oslo Breast Cancer Research Consortium, Oslo University Hospital, Oslo, Norway. [48]Department of Medical Genetics, Oslo University Hospital and University of Oslo, Oslo, Norway. [49]Department of Laboratory Medicine and Pathology, Mayo Clinic, Rochester, MN, USA. [50]Division of Clinical Medicine, School of Medicine and Population Health, University of Sheffield, Sheffield, UK. [51]Division of Neuroscience, School of Medicine and Population Health, University of Sheffield, Sheffield, UK. [52]Department of Medical Epidemiology and Biostatistics, Karolinska Institutet, Stockholm, Sweden. [53]Department of Clinical Genetics, Fox Chase Cancer Center, Philadelphia, PA, USA. [54]Department of Pathology, Leiden University Medical Center, Leiden, The Netherlands. [55]Department of Human Genetics, Leiden University Medical Center, Leiden, The Netherlands. [56]Department of Oncology, Centre for Cancer Genetic Epidemiology, University of Cambridge, Cambridge, UK. [57]Faculty of Medicine, University of Southampton, Southampton, UK. [58]Channing Division of Network Medicine, Department of Medicine, Brigham and Women's Hospital and Harvard Medical School, Boston, MA, USA. [59]Department of Nutrition, Harvard TH Chan School of Public Health, Boston, MA, USA. [60]Institute for Medical Informatics, Statistics and Epidemiology, University of Leipzig, Leipzig, Germany. [61]LIFE - Leipzig Research Centre for Civilization Diseases, University of Leipzig, Leipzig, Germany. [62]Division of Evolution and Genomic Sciences, School of Biological Sciences, Faculty of Biology, Medicine and Health, University of Manchester, Manchester Academic Health Science Centre, Manchester, UK. [63]North West Genomics Laboratory Hub, Manchester Centre for Genomic Medicine, St Mary's Hospital, Manchester University NHS Foundation Trust, Manchester Academic Health Science Centre, Manchester, UK. [64]Department of Gynecology and Obstetrics, Comprehensive Cancer Center Erlangen-EMN, Friedrich-Alexander University Erlangen-Nuremberg, University Hospital Erlangen, Erlangen, Germany. [65]The Breast Cancer Now Toby Robins Research Centre, The Institute of Cancer Research, London, UK. [66]Department of Breast Surgery, Herlev and Gentofte Hospital, Copenhagen University Hospital, Herlev, Denmark. [67]School of Population Health, Curtin University, Perth, WA, Australia. [68]Cancer Genetics and Epidemiology Group, Genomic Medicine Group, Fundación Instituto de Investigación Sanitaria de Santiago de Compostela (FIDIS), Complejo Hospitalario Universitario de Santiago, SERGAS, Santiago de Compostela, Spain. [69]MRC Clinical Trials Unit, Institute of Clinical Trials and Methodology, University College London, London, UK. [70]Department of Women's Cancer, Elizabeth Garrett Anderson Institute for Women's Health, University College London, London, UK. [71]Human Genotyping Unit-CeGen, Spanish National Cancer Research Centre (CNIO), Madrid, Spain. [72]Spanish Network on Rare Diseases (CIBERER), Madrid, Spain. [73]Team 'Exposome and Heredity', CESP, Gustave Roussy, INSERM, University Paris-Saclay, UVSQ, Villejuif, France. [74]Center for Familial Breast and Ovarian Cancer, Faculty of Medicine and University Hospital Cologne, University of Cologne, Cologne, Germany. [75]Center for Integrated Oncology (CIO), Faculty of Medicine and University Hospital Cologne, University of Cologne, Cologne, Germany. [76]Department of Preventive Medicine, Keck School of Medicine, University of Southern California, Los Angeles, CA, USA. [77]Molecular Genetics of Breast Cancer, German Cancer Research Center (DKFZ), Heidelberg, Germany. [78]Cancer RC, University of Eastern Finland, Kuopio, Finland. [79]Institute of Clinical Medicine, Pathology and Forensic Medicine, University of Eastern Finland, Kuopio, Finland. [80]Health Innovation and Evaluation Hub, Université de Montréal Hospital Research Centre (CRCHUM), Montréal, QC, Canada. [81]Department of Medical Oncology, Erasmus MC Cancer Institute, 3015 GD Rotterdam, The Netherlands. [82]Department of Gynecology and Obstetrics, University Hospital Düsseldorf, Heinrich-Heine University Düsseldorf, Düsseldorf, Germany. [83]University of Tübingen, Tübingen, Germany. [84]Centre for Epidemiology and Biostatistics, Melbourne School of Population and Global Health, The University of Melbourne, Melbourne, VIC, Australia. [85]Division of Cancer Sciences, Faculty of Biology, Medicine and Health, University of Manchester, Manchester Academic Health Science Centre, Manchester, UK. [86]Nightingale/Prevent Breast Cancer Centre, Wythenshawe Hospital, Manchester University NHS Foundation Trust, Manchester, UK. [87]Manchester Breast Centre, Manchester Cancer Research Centre, The Christie Hospital, Manchester, UK. [88]Division of Cancer Sciences, University of Manchester, Manchester, UK. [89]Australian Breast Cancer Tissue Bank, Westmead Institute for Medical Research, University of Sydney, Sydney, NSW, Australia. [90]Research Department, Peter MacCallum Cancer Center, Melbourne, VIC, Australia. [91]Sir Peter MacCallum Department of Oncology, The University of Melbourne, Parkville, VIC, Australia. [92]Research Centre for Genetic Engineering and Biotechnology 'Georgi D. Efremov', MASA, Skopje, Republic of North Macedonia. [93]Independent Laboratory of Molecular Biology and Genetic Diagnostics, Pomeranian Medical University, Szczecin, Poland. [94]Department of Genetics and Fundamental Medicine, Ufa University of Science and Technology, Ufa, Russia. [95]Division of Cancer Epidemiology and Genetics, National Cancer Institute, NIH, Bethesda, MD, USA. [96]Department of Computational and Quantitative Medicine, City of Hope, Duarte, CA, USA. [97]City of Hope Comprehensive Cancer Center, City of Hope, Duarte, CA, USA. [98]Laboratory for Translational Genetics, Department of Human Genetics, KU Leuven, Leuven, Belgium. [99]VIB Center for Cancer Biology, VIB, Leuven, Belgium. [100]Carmel Medical Center, Haifa, Israel. [101]Department of Molecular Medicine and Surgery, Karolinska Institutet, Stockholm, Sweden. [102]Department of Clinical Genetics, Karolinska University Hospital, Stockholm, Sweden. [103]Translational Cancer Research Area, University of Eastern Finland, Kuopio, Finland. [104]Biobank of Eastern Finland, Kuopio University Hospital, Kuopio, Finland. [105]Department of Medical Oncology, University Hospital of Heraklion, Heraklion, Greece. [106]School of Population and Public Health, University of British Columbia, Vancouver, BC, Canada. [107]Cancer Control Research, BC Cancer Agency, Vancouver, BC, Canada. [108]Department of Obstetrics and Gynecology, Helsinki University Hospital, University of Helsinki, Helsinki, Finland. [109]Institute for Occupational and Maritime Medicine, University Medical Center Hamburg-Eppendorf, Hamburg, Germany. [110]Institute for Medical Biometry and Epidemiology, University Medical Center Hamburg-Eppendorf, Hamburg, Germany. [111]Clinical Genetics Research Lab, Department of Cancer Biology and Genetics, Memorial Sloan Kettering Cancer Center, New York, NY, USA. [112]Clinical Genetics Service, Department of Medicine, Memorial Sloan Kettering Cancer Center, New York, NY, USA. [113]Genome Diagnostics Program, IFOM ETS - The AIRC Institute of Molecular Oncology, Milan, Italy. [114]Laboratory of Cancer Genetics and Tumor Biology, Translational Medicine Research Unit, Biocenter Oulu, University of Oulu, Oulu, Finland. [115]Laboratory of Cancer Genetics and Tumor Biology, Northern Finland Laboratory Centre Oulu, Oulu, Finland. [116]Unit of Predictive Medicine, Molecular Bases of Genetic Risk, Department of Experimental Oncology, Fondazione IRCCS Istituto Nazionale Dei Tumori (INT), Milan, Italy. [117]Department of Basic Sciences, Shaukat Khanum Memorial Cancer Hospital and Research Centre (SKMCH & RC), Lahore, Pakistan. [118]Technion, Faculty of Medicine and Association for Promotion of Research in Precision Medicine, Haifa, Israel. [119]Medical Oncology Department, Hospital

Universitario Puerta de Hierro, Madrid, Spain. [120]Department of Pathology, The Netherlands Cancer Institute - Antoni Van Leeuwenhoek Hospital, Amsterdam, The Netherlands. [121]Department of Oncology, University Hospital of Larissa, Larissa, Greece. [122]Epidemiology Branch, National Institute of Environmental Health Sciences, NIH, Research Triangle Park, NC, USA. [123]School of Cancer and Pharmaceutical Sciences, Comprehensive Cancer Centre, Guy's Campus, King's College London, London, UK. [124]Center for Molecular Medicine Cologne (CMMC), Faculty of Medicine and University Hospital Cologne, University of Cologne, Cologne, Germany. [125]Division of Epidemiology, Department of Medicine, Vanderbilt Epidemiology Center, Vanderbilt-Ingram Cancer Center, Vanderbilt University School of Medicine, Nashville, TN, USA. [126]Precision Medicine, School of Clinical Sciences at Monash Health, Monash University, Clayton, VIC, Australia. [127]Department of Clinical Pathology, The University of Melbourne, Melbourne, VIC, Australia. [128]Cancer Epidemiology Division, Cancer Council Victoria, Melbourne, VIC, Australia. [129]Genetic Epidemiology Group, School of Population and Global Health, University of Western Australia, Perth, WA, Australia. [130]Epigenetic and Stem Cell Biology Laboratory, National Institute of Environmental Health Sciences, NIH, Research Triangle Park, NC, USA. [131]Department of Clinical Genetics, The Netherlands Cancer Institute - Antoni Van Leeuwenhoek Hospital, Amsterdam, The Netherlands. [132]Division of Epidemiology, Department of Quantitative Health Sciences, Mayo Clinic, Rochester, MN, USA. [133]Division of Molecular Pathology, The Netherlands Cancer Institute, Amsterdam, The Netherlands. [134]Division of Psychosocial Research and Epidemiology, The Netherlands Cancer Institute - Antoni Van Leeuwenhoek Hospital, Amsterdam, The Netherlands. [135]Department of Clinical Genetics, Leiden University Medical Center, Leiden, The Netherlands. [136]Department of Oncology, Södersjukhuset, Stockholm, Sweden. [137]Department of Computational Biomedicine, Cedars-Sinai Medical Center, West Hollywood, CA, USA. [138]Genomics Center, Centre Hospitalier Universitaire de Québec – Université Laval Research Center, Québec City, QC, Canada.

Received: 12 August 2024 Accepted: 9 December 2024
Published online: 29 December 2024

References

1. Michailidou K, Hall P, Gonzalez-Neira A, et al. Large-scale genotyping identifies 41 new loci associated with breast cancer risk. Nat Genet. 2013;45(4):353–61.
2. Michailidou K, Beesley J, Lindstrom S, et al. Genome-wide association analysis of more than 120,000 individuals identifies 15 new susceptibility loci for breast cancer. Nat Genet. 2015;47(4):373–80.
3. Michailidou K, Lindström S, Dennis J, et al. Association analysis identifies 65 new breast cancer risk loci. Nature. 2017;551(7678):92–4.
4. Zhang H, Ahearn TU, Lecarpentier J, et al. Genome-wide association study identifies 32 novel breast cancer susceptibility loci from overall and subtype-specific analyses. Nat Genet. 2020;52(6):572–81.
5. Dorling L, Carvalho S, Allen J, et al. Breast cancer risk genes—association analysis in more than 113,000 women. N Engl J Med. 2021;384(5):428–39.
6. Kuchenbaecker KB, Hopper JL, Barnes DR, et al. Risks of breast, ovarian, and contralateral breast cancer for BRCA1 and BRCA2 mutation carriers. JAMA. 2017;317(23):2402–16.
7. Choi SW, Mak TS, O'Reilly PF. Tutorial: a guide to performing polygenic risk score analyses. Nat Protoc. 2020;15(9):2759–72.
8. Wand H, Lambert SA, Tamburro C, et al. Improving reporting standards for polygenic scores in risk prediction studies. Nature. 2021;591(7849):211–9.
9. Mavaddat N, Pharoah PD, Michailidou K, et al. Prediction of breast cancer risk based on profiling with common genetic variants. J Natl Cancer Inst. 2015;107(5):djv036.
10. Khera AV, Chaffin M, Aragam KG, et al. Genome-wide polygenic scores for common diseases identify individuals with risk equivalent to monogenic mutations. Nat Genet. 2018;50(9):1219–24.
11. Mavaddat N, Michailidou K, Dennis J, et al. Polygenic risk scores for prediction of breast cancer and breast cancer subtypes. Am J Hum Genet. 2019;104(1):21–34.
12. Shieh Y, Eklund M, Madlensky L, et al. Breast cancer screening in the precision medicine era: risk-based screening in a population-based trial. J Natl Cancer Inst. 2017;109(5):djw290.
13. Pashayan N, Morris S, Gilbert FJ, Pharoah PDP. Cost-effectiveness and benefit-to-harm ratio of risk-stratified screening for breast cancer: a life-table model. JAMA Oncol. 2018;4(11):1504–10.
14. 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):1708–18.
15. Lewis CM, Vassos E. Polygenic risk scores: from research tools to clinical instruments. Genome Med. 2020;12(1):44.
16. Pashayan N, Antoniou AC, Ivanus U, et al. Personalized early detection and prevention of breast cancer: ENVISION consensus statement. Nat Rev Clin Oncol. 2020;17(11):687–705.
17. Brooks JD, Nabi HH, Andrulis IL, et al. Personalized risk assessment for prevention and early detection of breast cancer: integration and implementation (PERSPECTIVE I&I). J Pers Med. 2021;11(6):511.
18. van den Broek JJ, Schechter CB, van Ravesteyn NT, et al. Personalizing breast cancer screening based on polygenic risk and family history. J Natl Cancer Inst. 2021;113(4):434–42.
19. Pashayan N, Easton DF, Michailidou K. Polygenic risk scores in cancer screening: A glass half full or half empty? Lancet Oncol. 2023;24(6):579–81.
20. Yang X, Kar S, Antoniou AC, Pharoah PDP. Polygenic scores in cancer. Nat Rev Cancer. 2023;23(9):619–30.
21. Carver T, Hartley S, Lee A, et al. CanRisk tool—a web interface for the prediction of breast and ovarian cancer risk and the likelihood of carrying genetic pathogenic variants. Cancer Epidemiol Biomark Prev. 2021;30(3):469–73.
22. Archer S, Babb de Villiers C, Scheibl F, et al. Evaluating clinician acceptability of the prototype CanRisk tool for predicting risk of breast and ovarian cancer: A multi-methods study. PLoS One. 2020;15(3):e0229999.
23. Lakeman IMM, Rodríguez-Girondo M, Lee A, et al. Validation of the BOADICEA model and a 313-variant polygenic risk score for breast cancer risk prediction in a Dutch prospective cohort. Genet Med. 2020;22(11):1803–11.
24. Pal Choudhury P, Brook MN, Hurson AN, et al. Comparative validation of the BOADICEA and Tyrer-Cuzick breast cancer risk models incorporating classical risk factors and polygenic risk in a population-based prospective cohort of women of European ancestry. Breast Cancer Res. 2021;23(1):22.
25. Li SX, Milne RL, Nguyen-Dumont T, et al. Prospective evaluation of the addition of polygenic risk scores to breast cancer risk models. JNCI Cancer Spectr. 2021;5(3):pkab021.
26. Yang X, Eriksson M, Czene K, et al. Prospective validation of the BOADICEA multifactorial breast cancer risk prediction model in a large prospective cohort study. J Med Genet. 2022;59(12):1196–205.
27. Lee A, Mavaddat N, Cunningham A, 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. 2022;59(12):1206–18.
28. Ho WK, Tan MM, Mavaddat N, et al. European polygenic risk score for prediction of breast cancer shows similar performance in Asian women. Nat Commun. 2020;11(1):3833.
29. Du Z, Gao G, Adedokun B, et al. Evaluating polygenic risk scores for breast cancer in women of African ancestry. J Natl Cancer Inst. 2021;113(9):1168–76.
30. Liu C, Zeinomar N, Chung WK, et al. Generalizability of polygenic risk scores for breast cancer among women with European, African, and Latinx Ancestry. JAMA Netw Open. 2021;4(8):e2119084.
31. Amos CI, Dennis J, Wang Z, et al. The OncoArray consortium: a network for understanding the genetic architecture of common cancers. Cancer Epidemiol Biomark Prev. 2017;26(1):126–35.
32. Li Y, Byun J, Cai G, et al. FastPop: a rapid principal component derived method to infer intercontinental ancestry using genetic data. BMC Bioinform. 2016;17:122.
33. Bycroft C, Freeman C, Petkova D, et al. The UK Biobank resource with deep phenotyping and genomic data. Nature. 2018;562(7726):203–9.
34. 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.
35. Thompson DJ, Wells D, Selzam S, et al. UK Biobank release and systematic evaluation of optimised polygenic risk scores for 53 diseases and quantitative traits. medRxiv. https://doi.org/10.1101/2022.06.16.22276246, 16 June 2022, preprint: not peer reviewed.

36. Schmidt MK, Hogervorst F, van Hien R, et al. Age- and tumor subtype-specific breast cancer risk estimates for CHEK2*1100delC carriers. J Clin Oncol. 2016;34(23):2750–60.
37. Clayton D, Kaldor J. Empirical Bayes estimates of age-standardized relative risks for use in disease mapping. Biometrics. 1987;43(3):671–81.
38. Wang Y, Tsuo K, Kanai M, Neale BM, Martin AR. Challenges and opportunities for developing more generalizable polygenic risk scores. Annu Rev Biomed Data Sci. 2022;5:293–320.
39. Martin AR, Kanai M, Kamatani Y, Okada Y, Neale BM, Daly MJ. Clinical use of current polygenic risk scores may exacerbate health disparities. Nat Genet. 2019;51(4):584–91.
40. National Institute for Health and Care Excellence: Guidelines. In: Familial breast cancer: classification, care and managing breast cancer and related risks in people with a family history of breast cancer. London: National Institute for Health and Care Excellence (NICE) Copyright © NICE 2020.; 2019.
41. Ho WK, Tai MC, Dennis J, et al. Polygenic risk scores for prediction of breast cancer risk in Asian populations. Genet Med. 2022;24(3):586–600.
42. Sung H, Ferlay J, Siegel RL, et al. Global cancer statistics 2020: GLOBOCAN estimates of incidence and mortality worldwide for 36 cancers in 185 countries. CA Cancer J Clin. 2021;71(3):209–49.

Publisher's Note