Breast cancer polygenic risk score and contralateral breast cancer risk

Iris Kramer[1], Maartje J. Hooning[2], Nasim Mavaddat[3], Michael Hauptmann[4, 5], Renske Keeman[1], Ewout W Steyerberg[6, 7], Daniele Giardiello[1, 6], Antonis C. Antoniou[3], Paul D.P. Pharoah[3, 8], Sander Canisius[1, 9], Zumuruda Abu-Ful [10], Irene L. Andrulis[11, 12], Hoda Anton-Culver[13], Kristan J. Aronson[14], Annelie Augustinsson[15], Heiko Becher[16, 17], Matthias W. Beckmann[18], Sabine Behrens[19], Javier Benitez[20, 21], Marina Bermisheva[22], Natalia V. Bogdanova[23-25], Stig E. Bojesen[26-28], Manjeet K. Bolla[3], Bernardo Bonanni[29], Hiltrud Brauch[30-32], Michael Bremer[23], Sara Y. Brucker[33], Barbara Burwinkel[34, 35], Jose E. Castelao[36], Tsun L. Chan[37, 38], Jenny Chang-Claude[19, 39], Stephen J. Chanock[40], Georgia Chenevix-Trench[41], Ji-Yeob Choi[42, 43], Christine L. Clarke[44], NBCS Collaborators[45-56], J. Margriet Collée[57], Fergus J. Couch[58], Angela Cox[59], Simon S. Cross[60], Kamila Czene[61], Mary B. Daly[62], Peter Devilee[63, 64], Thilo Dörk[24], Isabel dos-Santos-Silva[65], Alison M. Dunning[8], Miriam Dwek[66], Diana M. Eccles[67], D. Gareth Evans[68, 69], Peter A. Fasching[18, 70], Henrik Flyger[71], Manuela Gago-Dominguez[72, 73], Montserrat García-Closas[40], José A. García-Sáenz[74], Graham G. Giles[75-77], David E. Goldgar[78], Anna González-Neira[21], Christopher A. Haiman[79], Niclas Håkansson[80], Ute Hamann[81], Mikael Hartman[82, 83], Bernadette A.M. Heemskerk-Gerritsen[2], Antoinette Hollestelle[2], John L. Hopper[76], Ming-Feng Hou[84], Anthony Howell[85], ABCTB Investigators[86], kConFab Investigators[87, 88], Hidemi Ito[89, 90], Milena Jakimovska[91], Anna Jakubowska[92, 93], Wolfgang Janni[94], Esther M. John[95], Audrey Jung[19], Daehee Kang[42, 43, 96], C. Marleen Kets[97], Elza Khusnutdinova[22, 98], Yon-Dschun Ko[99], Vessela N. Kristensen[45, 56], Allison W. Kurian[95, 100], Ava Kwong[37, 101, 102], Diether Lambrechts[103, 104], Loic Le Marchand[105], Jingmei Li[106], Annika Lindblom[107, 108], Jan Lubiński[92], Arto Mannermaa[109-111], Mehdi Manoochehri[81], Sara Margolin[112, 113], Keitaro Matsuo[89, 90], Dimitrios Mavroudis[114], Alfons Meindl[115], Roger L. Milne[75-77], Anna Marie Mulligan[116, 117], Taru A. Muranen[118], Susan L. Neuhausen[119], Heli Nevanlinna[118], William G. Newman[68, 69], Andrew F. Olshan[120], Janet E.

Olson[121], Håkan Olsson[15], Tjoung-Won Park-Simon[24], Julian Peto[65], Christos Petridis[122], Dijana Plaseska-Karanfilska[91], Nadege Presneau[66], Katri Pylkäs[123, 124], Paolo Radice[125], Gad Rennert[10], Atocha Romero[126], Rebecca Roylance[127], Emmanouil Saloustros[128], Elinor J. Sawyer[129], Rita K. Schmutzler[130-132], Lukas Schwentner[94], Christopher Scott[121], Mee-Hoong See[133], Mitul Shah[8], Chen-Yang Shen[134, 135], Xiao-Ou Shu[136], Sabine Siesling[137, 138], Susan Slager[121], Christof Sohn[139], Melissa C. Southey[75, 77, 140], John J. Spinelli[141, 142], Jennifer Stone[76, 143], William J. Tapper[67], Maria Tengström[109, 144, 145], Soo Hwang Teo[146, 147], Mary Beth Terry[148], Rob A.E.M. Tollenaar[149], Ian Tomlinson[150, 151], Melissa A. Troester[120], Celine M. Vachon[152], Chantal van Ongeval[153], Elke M. van Veen[68, 69], Robert Winqvist[123, 124], Alicja Wolk[80, 154], Wei Zheng[136], Argyrios Ziogas[13], Douglas F. Easton[3, 8], Per Hall[61, 112], Marjanka K. Schmidt[1, 155]

[1] The Netherlands Cancer Institute - Antoni van Leeuwenhoek Hospital, Division of Molecular Pathology, Amsterdam, 1066 CX, The Netherlands.

[2] Erasmus MC Cancer Institute, Department of Medical Oncology, Rotterdam, 3015 CN, The Netherlands.

[3] University of Cambridge, Centre for Cancer Genetic Epidemiology, Department of Public Health and Primary Care, Cambridge, CB1 8RN, UK.

[4] The Netherlands Cancer Institute - Antoni van Leeuwenhoek hospital, Department of Epidemiology and Biostatistics, Amsterdam, 1066 CX, The Netherlands.

[5] Brandenburg Medical School Theodor Fontane, Institute of Biostatistics and Registry Research, Neuruppin, 16816, Germany.

[6] Leiden University Medical Center, Department of Biomedical Data Sciences, Leiden, 2333 ZA, The Netherlands.

[7] Erasmus MC Cancer Institute, Department of Public Health, Rotterdam, 3015 GD, The Netherlands.

[8] University of Cambridge, Centre for Cancer Genetic Epidemiology, Department of Oncology, Cambridge, CB1 8RN, UK.

[9] The Netherlands Cancer Institute - Antoni van Leeuwenhoek hospital, Division of Molecular Carcinogenesis, Amsterdam, 1066 CX, The Netherlands.

[10] Carmel Medical Center and Technion Faculty of Medicine, Clalit National Cancer Control Center, Haifa, 35254, Israel.

[11] Lunenfeld-Tanenbaum Research Institute of Mount Sinai Hospital, Fred A. Litwin Center for Cancer Genetics, Toronto, ON, M5G 1X5, Canada.

[12] University of Toronto, Department of Molecular Genetics, Toronto, ON, M5S 1A8, Canada.

[13] University of California Irvine, Department of Epidemiology, Genetic Epidemiology Research Institute, Irvine, CA, 92617, USA.

[14] Queen's University, Department of Public Health Sciences, and Cancer Research Institute, Kingston, ON, K7L 3N6, Canada.

[15] Lund University, Department of Cancer Epidemiology, Clinical Sciences, Lund, 222 42, Sweden.

[16] University Medical Center Hamburg-Eppendorf, Institute of Medical Biometry and Epidemiology, Hamburg, 20246, Germany.

[17] Charité –Universitätsmedizin Berlin, Institute of Biometry and Clinical Epidemiology, Berlin, 10117, Germany.

[18] University Hospital Erlangen, Friedrich-Alexander-University Erlangen-Nuremberg, Department of Gynecology and Obstetrics, Comprehensive Cancer Center ER-EMN, Erlangen, 91054, Germany.

[19] German Cancer Research Center (DKFZ), Division of Cancer Epidemiology, Heidelberg, 69120, Germany.

[20] Centro de Investigación en Red de Enfermedades Raras (CIBERER), Madrid, 28029, Spain.

[21] Spanish National Cancer Research Centre (CNIO), Human Cancer Genetics Programme, Madrid, 28029, Spain.

[22] Ufa Federal Research Centre of the Russian Academy of Sciences, Institute of Biochemistry and Genetics, Ufa, 450054, Russia.

[23] Hannover Medical School, Department of Radiation Oncology, Hannover, 30625, Germany.

[24] Hannover Medical School, Gynaecology Research Unit, Hannover, 30625, Germany.

[25] N.N. Alexandrov Research Institute of Oncology and Medical Radiology, Minsk, 223040, Belarus.

[26] Copenhagen University Hospital, Copenhagen General Population Study, Herlev and Gentofte Hospital, Herlev, 2730, Denmark.

[27] Copenhagen University Hospital, Department of Clinical Biochemistry, Herlev and Gentofte Hospital, Herlev, 2730, Denmark.

[28] University of Copenhagen, Faculty of Health and Medical Sciences, Copenhagen, 2200, Denmark.

[29] IEO, European Institute of Oncology IRCCS, Division of Cancer Prevention and Genetics, Milan, 20141, Italy.

[30] Dr. Margarete Fischer-Bosch-Institute of Clinical Pharmacology, Stuttgart, 70376, Germany.

[31] University of Tübingen, iFIT-Cluster of Excellence, Tübingen, 72074, Germany.

[32] German Cancer Research Center (DKFZ), German Cancer Consortium (DKTK), Partner Site Tübingen, Tübingen, 72074, Germany.

[33] University of Tübingen, Department of Gynecology and Obstetrics, Tübingen, 72076, Germany.

[34] German Cancer Research Center (DKFZ), Molecular Epidemiology Group, C080, Heidelberg, 69120, Germany.

[35] University of Heidelberg, Molecular Biology of Breast Cancer, University Womens Clinic Heidelberg, Heidelberg, 69120, Germany.

[36] Instituto de Investigacion Sanitaria Galicia Sur (IISGS), Xerencia de Xestion Integrada de Vigo-SERGAS, Oncology and Genetics Unit, Vigo, 36312, Spain.

[37] Cancer Genetics Centre, Hong Kong Hereditary Breast Cancer Family Registry, Happy Valley, Hong Kong.

[38] Hong Kong Sanatorium and Hospital, Department of Pathology, Happy Valley, Hong Kong.

[39] University Medical Center Hamburg-Eppendorf, Cancer Epidemiology Group, University Cancer Center Hamburg (UCCH), Hamburg, 20246, Germany.

[40] National Cancer Institute, National Institutes of Health, Department of Health and Human Services, Division of Cancer Epidemiology and Genetics, Bethesda, MD, 20850, USA.

[41] QIMR Berghofer Medical Research Institute, Department of Genetics and Computational Biology, Brisbane, Queensland, 4006, Australia.

[42] Seoul National University Graduate School, Department of Biomedical Sciences, Seoul, 03080, Korea.

[43] Seoul National University, Cancer Research Institute, Seoul, 03080, Korea.

[44] University of Sydney, Westmead Institute for Medical Research, Sydney, New South Wales, 2145, Australia.

[45] Oslo University Hospital-Radiumhospitalet, Department of Cancer Genetics, Institute for Cancer Research, Oslo, 0379, Norway.

[46] University of Oslo, Institute of Clinical Medicine, Faculty of Medicine, Oslo, 0450, Norway.

[47] Vestre Viken Hospital, Department of Research, Drammen, 3019, Norway.

[48] Oslo University Hospital-Ullevål, Section for Breast- and Endocrine Surgery, Department of Cancer, Division of Surgery, Cancer and Transplantation Medicine, Oslo, 0450, Norway.

[49] Oslo University Hospital, Department of Radiology and Nuclear Medicine, Oslo, 0379, Norway.

[50] Akershus University Hospital, Department of Pathology, Lørenskog, 1478, Norway.

[51] Oslo University Hospital, Department of Tumor Biology, Institute for Cancer Research, Oslo, 0379, Norway.

[52] Oslo University Hospital-Radiumhospitalet, Department of Oncology, Division of Surgery, Cancer and Transplantation Medicine, Oslo, 0379, Norway.

[53] Oslo University Hospital-Radiumhospitalet, National Advisory Unit on Late Effects after Cancer Treatment, Department of Oncology, Oslo, 0379, Norway.

[54] Akershus University Hospital, Department of Oncology, Lørenskog, 1478, Norway.

[55] Oslo University Hospital, Oslo Breast Cancer Research Consortium, Oslo, 0379, Norway.

[56] Oslo University Hospital and University of Olso, Department of Medical Genetics, Oslo, 0379, Norway.

[57] Erasmus University Medical Center, Department of Clinical Genetics, Rotterdam, 3015 CN, The Netherlands.

[58] Mayo Clinic, Department of Laboratory Medicine and Pathology, Rochester, MN, 55905, USA.

[59] University of Sheffield, Sheffield Institute for Nucleic Acids (SInFoNiA), Department of Oncology and Metabolism, Sheffield, S10 2TN, UK.

[60] University of Sheffield, Academic Unit of Pathology, Department of Neuroscience, Sheffield, S10 2TN, UK.

[61] Karolinska Institutet, Department of Medical Epidemiology and Biostatistics, Stockholm, 171 65, Sweden.

[62] Fox Chase Cancer Center, Department of Clinical Genetics, Philadelphia, PA, 19111, USA.

[63] Leiden University Medical Center, Department of Pathology, Leiden, 2333 ZA, The Netherlands.

[64] Leiden University Medical Center, Department of Human Genetics, Leiden, 2333 ZA, The Netherlands.

[65] London School of Hygiene and Tropical Medicine, Department of Non-Communicable Disease Epidemiology, London, WC1E 7HT, UK.

[66] University of Westminster, School of Life Sciences, London, W1B 2HW, UK.

[67] University of Southampton, Faculty of Medicine, Southampton, SO17 1BJ, UK.

[68] University of Manchester, Manchester Academic Health Science Centre, Division of Evolution and Genomic Sciences, School of Biological Sciences, Faculty of Biology, Medicine and Health, Manchester, M13 9WL, UK.

[69] St Mary's Hospital, Manchester University NHS Foundation Trust, Manchester Academic Health Science Centre, North West Genomics Laboratory Hub, Manchester Centre for Genomic Medicine, Manchester, M13 9WL, UK.

[70] University of California at Los Angeles, David Geffen School of Medicine, Department of Medicine Division of Hematology and Oncology, Los Angeles, CA, 90095, USA.

[71] Copenhagen University Hospital, Department of Breast Surgery, Herlev and Gentofte Hospital, Herlev, 2730, Denmark.

[72] Grupo de Medicina Xenómica, Instituto de Investigación Sanitaria de Santiago de Compostela (IDIS), Fundación Pública Galega de Medicina Xenómica, Santiago de Compostela, 15706, Spain.

[73] University of California San Diego, Moores Cancer Center, La Jolla, CA, 92037, USA.

[74] Instituto de Investigación Sanitaria San Carlos (IdISSC), Centro Investigación Biomédica en Red de Cáncer (CIBERONC), Medical Oncology Department, Hospital Clínico San Carlos, Madrid, 28040, Spain.

[75] Cancer Council Victoria, Cancer Epidemiology Division, Melbourne, Victoria, 3004, Australia.

[76] The University of Melbourne, Centre for Epidemiology and Biostatistics, Melbourne School of Population and Global Health, Melbourne, Victoria, 3010, Australia.

[77] Monash University, Precision Medicine, School of Clinical Sciences at Monash Health, Clayton, Victoria, 3168, Australia.

[78] Huntsman Cancer Institute, University of Utah School of Medicine, Department of Dermatology, Salt Lake City, UT, 84112, USA.

[79] University of Southern California, Department of Preventive Medicine, Keck School of Medicine, Los Angeles, CA, 90033, USA.

[80] Karolinska Institutet, Institute of Environmental Medicine, Stockholm, 171 77, Sweden.

[81] German Cancer Research Center (DKFZ), Molecular Genetics of Breast Cancer, Heidelberg, 69120, Germany.

[82] National University of Singapore and National University Health System, Saw Swee Hock School of Public Health, Singapore, 119077, Singapore.

[83] National University Health System, Department of Surgery, Singapore, 119228, Singapore.

[84] Kaohsiung Medical University, Chung-Ho Memorial Hospital, Kaohsiung, 807, Taiwan.

[85] University of Manchester, Division of Cancer Sciences, Manchester, M13 9PL, UK.

[86] University of Sydney, Australian Breast Cancer Tissue Bank, Westmead Institute for Medical Research, Sydney, New South Wales, 2145, Australia.

[87] Peter MacCallum Cancer Center, Melbourne, Victoria, 3000, Australia.

[88] The University of Melbourne, Sir Peter MacCallum Department of Oncology, Melbourne, Victoria, 3000, Australia.

[89] Aichi Cancer Center Research Institute, Division of Cancer Epidemiology and Prevention, Nagoya, 464-8681, Japan.

[90] Nagoya University Graduate School of Medicine, Division of Cancer Epidemiology, Nagoya, 466-8550, Japan.

[91] MASA, Research Centre for Genetic Engineering and Biotechnology 'Georgi D. Efremov', Skopje, 1000, Republic of North Macedonia.

[92] Pomeranian Medical University, Department of Genetics and Pathology, Szczecin, 71-252, Poland.

[93] Pomeranian Medical University, Independent Laboratory of Molecular Biology and Genetic Diagnostics, Szczecin, 71-252, Poland.

[94] University Hospital Ulm, Department of Gynaecology and Obstetrics, Ulm, 89075, Germany.

[95] Stanford Cancer Institute, Stanford University School of Medicine, Department of Epidemiology & Population Health, Stanford, CA, 94304, USA.

[96] Seoul National University College of Medicine, Department of Preventive Medicine, Seoul, 03080, Korea.

[97] The Netherlands Cancer Institute - Antoni van Leeuwenhoek Hospital, Department of Clinical Genetics, Amsterdam, 1066 CX, The Netherlands.

[98] Bashkir State University, Department of Genetics and Fundamental Medicine, Ufa, 450000, Russia.

[99] Johanniter Krankenhaus, Department of Internal Medicine, Evangelische Kliniken Bonn gGmbH, Bonn, 53177, Germany.

[100] Stanford University School of Medicine, Department of Health Research and Policy, Stanford, CA, 94305, USA.

[101] The University of Hong Kong, Department of Surgery, Pok Fu Lam, Hong Kong.

[102] Hong Kong Sanatorium and Hospital, Cancer Genetics Center and Department of Surgery, Happy Valley, Hong Kong.

[103] VIB Center for Cancer Biology, Leuven, 3001, Belgium.

[104] University of Leuven, Laboratory for Translational Genetics, Department of Human Genetics, Leuven, 3000, Belgium.

[105] University of Hawaii Cancer Center, Epidemiology Program, Honolulu, HI, 96813, USA.

[106] Genome Institute of Singapore, Human Genetics Division, Singapore, 138672, Singapore.

[107] Karolinska Institutet, Department of Molecular Medicine and Surgery, Stockholm, 171 76, Sweden.

[108] Karolinska University Hospital, Department of Clinical Genetics, Stockholm, 171 76, Sweden.

[109] University of Eastern Finland, Translational Cancer Research Area, Kuopio, 70210, Finland.

[110] University of Eastern Finland, Institute of Clinical Medicine, Pathology and Forensic Medicine, Kuopio, 70210, Finland.

[111] Kuopio University Hospital, Biobank of Eastern Finland, Kuopio, 70210, Finland.

[112] Södersjukhuset, Department of Oncology, Stockholm, 118 83, Sweden.

[113] Karolinska Institutet, Department of Clinical Science and Education, Södersjukhuset, Stockholm, 118 83, Sweden.

[114] University Hospital of Heraklion, Department of Medical Oncology, Heraklion, 711 10, Greece.

[115] University of Munich, Campus Großhadern, Department of Gynecology and Obstetrics, Munich, 81377, Germany.

[116] University of Toronto, Department of Laboratory Medicine and Pathobiology, Toronto, ON, M5S 1A8, Canada.

[117] University Health Network, Laboratory Medicine Program, Toronto, ON, M5G 2C4, Canada.

[118] Helsinki University Hospital, Department of Obstetrics and Gynecology, University of Helsinki, Helsinki, 00290, Finland.

[119] Beckman Research Institute of City of Hope, Department of Population Sciences, Duarte, CA, 91010, USA.

[120] University of North Carolina at Chapel Hill, Department of Epidemiology, Gillings School of Global Public Health and UNC Lineberger Comprehensive Cancer Center, Chapel Hill, NC, 27599, USA.

[121] Mayo Clinic, Department of Health Sciences Research, Rochester, MN, 55905, USA.

[122] King's College London, Research Oncology, Guy's Hospital, London, SE1 9RT, UK.

[123] University of Oulu, Laboratory of Cancer Genetics and Tumor Biology, Cancer and Translational Medicine Research Unit, Biocenter Oulu, Oulu, 90220, Finland.

[124] Northern Finland Laboratory Centre Oulu, Laboratory of Cancer Genetics and Tumor Biology, Oulu, 90220, Finland.

[125] Fondazione IRCCS Istituto Nazionale dei Tumori (INT), Unit of Molecular Bases of Genetic Risk and Genetic Testing, Department of Research, Milan, 20133, Italy.

[126] Hospital Universitario Puerta de Hierro, Medical Oncology Department, Madrid, 28222, Spain.

[127] UCLH Foundation Trust, Department of Oncology, London, NW1 2PG, UK.

[128] University Hospital of Larissa, Department of Oncology, Larissa, 411 10, Greece.

[129] King's College London, School of Cancer & Pharmaceutical Sciences, Comprehensive Cancer Centre, Guy's Campus, London, SE1 1UL, UK.

[130] Faculty of Medicine and University Hospital Cologne, University of Cologne, Center for Familial Breast and Ovarian Cancer, Cologne, 50937, Germany.

[131] Faculty of Medicine and University Hospital Cologne, University of Cologne, Center for Integrated Oncology (CIO), Cologne, 50937, Germany.

[132] Faculty of Medicine and University Hospital Cologne, University of Cologne, Center for Molecular Medicine Cologne (CMMC), Cologne, 50931, Germany.

[133] University of Malaya, Breast Cancer Research Unit, University Malaya Cancer Research Institute, Faculty of Medicine, Kuala Lumpur, 50603, Malaysia.

[134] Academia Sinica, Institute of Biomedical Sciences, Taipei, 115, Taiwan.

[135] China Medical University, School of Public Health, Taichung, 40402, Taiwan.

[136] Vanderbilt University School of Medicine, Division of Epidemiology, Department of Medicine, Vanderbilt Epidemiology Center, Vanderbilt-Ingram Cancer Center, Nashville, TN, 37232, USA.

[137] Netherlands Comprehensive Cancer Organisation (IKNL), Department of Research, Utrecht, 3511 DT, The Netherlands.

[138] University of Twente, Department of Health Technology and Service Research, Technical Medical Center, Enschede, 7522 NB, The Netherlands.

[139] University Hospital and German Cancer Research Center, National Center for Tumor Diseases, Heidelberg, 69120, Germany.

[140] The University of Melbourne, Department of Clinical Pathology, Melbourne, Victoria, 3010, Australia.

[141] BC Cancer, Population Oncology, Vancouver, BC, V5Z 1G1, Canada.

[142] University of British Columbia, School of Population and Public Health, Vancouver, BC, V6T 1Z4, Canada.

[143] Curtin University and University of Western Australia, The Curtin UWA Centre for Genetic Origins of Health and Disease, Perth, Western Australia, 6000, Australia.

[144] Kuopio University Hospital, Department of Oncology, Cancer Center, Kuopio, 70210, Finland.

[145] University of Eastern Finland, Institute of Clinical Medicine, Oncology, Kuopio, 70210, Finland.

[146] Cancer Research Malaysia, Breast Cancer Research Programme, Subang Jaya, Selangor, 47500, Malaysia.

[147] University of Malaya, Department of Surgery, Faculty of Medicine, Kuala Lumpur, 50603, Malaysia.

[148] Columbia University, Department of Epidemiology, Mailman School of Public Health, New York, NY, 10032, USA.

[149] Leiden University Medical Center, Department of Surgery, Leiden, 2333 ZA, The Netherlands.

[150] University of Birmingham, Institute of Cancer and Genomic Sciences, Birmingham, B15 2TT, UK.

[151] University of Oxford, Wellcome Trust Centre for Human Genetics and Oxford NIHR Biomedical Research Centre, Oxford, OX3 7BN, UK.

[152] Mayo Clinic, Department of Health Science Research, Division of Epidemiology, Rochester, MN, 55905, USA.

[153] Leuven Cancer Institute, University Hospitals Leuven, Leuven Multidisciplinary Breast Center, Department of Radiology, Leuven, 3000, Belgium.

[154] Uppsala University, Department of Surgical Sciences, Uppsala, 751 05, Sweden.

[155] The Netherlands Cancer Institute - Antoni van Leeuwenhoek hospital, Division of Psychosocial Research and Epidemiology, Amsterdam, 1066 CX, The Netherlands.

***Correspondence:** Marjanka K Schmidt, Netherlands Cancer Institute, Plesmanlaan 121, 1066 CX Amsterdam, The Netherlands; +31205122767; mk.schmidt@nki.nl

Abstract

Previous research has shown that polygenic risk scores (PRS) can be used to stratify women according to their risk of developing primary invasive breast cancer. This study aimed to evaluate the association between a recently validated PRS of 313 germline variants (PRS_{313}) and contralateral breast cancer (CBC) risk. We included 56,068 women of European ancestry diagnosed with first invasive breast cancer from 1990 onwards with follow-up from the Breast Cancer Association Consortium. Metachronous CBC risk (N=1,027) according to the distribution of the PRS_{313} was quantified using Cox regression analyses. We assessed PRS_{313} interaction with age at first diagnosis, family history, morphology, ER-, PR-, and HER2-status, and (neo)adjuvant therapy. In Asian studies, with limited follow-up, CBC risk associated with PRS_{313} was assessed using logistic regression for 340 women with CBC compared with 12,133 women with unilateral breast cancer. Higher PRS_{313} was associated with increased CBC risk: hazard ratio per standard deviation (SD)=1.25 (95%CI=1.18-1.33) for Europeans, and an OR per SD=1.15 (95%CI=1.02-1.29) for Asians. The absolute lifetime risks of CBC, accounting for death as competing risk, were 12.4% for European women at the 10^{th} percentile and 20.5% at the 90^{th} percentile of the PRS_{313}. We found no evidence of confounding by, or interaction with patient characteristics, characteristics of the primary tumor, or treatment. The C-index for the PRS_{313} alone was 0.563 (95%CI=0.547-0.586). In conclusion, the PRS_{313} is an independent factor associated with CBC risk, and may be incorporated in CBC risk prediction models to help improve stratification of patients and optimize surveillance and treatment strategies.

Introduction

Due to the high incidence of breast cancer and improving survival, an increasing number of breast cancer survivors are at risk of developing contralateral breast cancer (CBC). The 10-year cumulative incidence of CBC is ~4%[1; 2], however estimates vary widely depending on factors such as germline genetics, family history, and (neo)adjuvant systemic therapy for the first breast cancer[3]. The risk of developing CBC is particularly high in women with rare mutations in certain genes including *BRCA1, BRCA2,* and *CHEK2*, with approximately two- to fourfold higher risks reported compared with women without these mutations[3].

Recently, genome-wide association studies (GWAS) have identified multiple common germline variants that are associated with first primary breast cancer risk[4; 5]. These are associated with small differences in risk individually, but their combined effects can be summarized in a polygenic risk score (PRS), which has been shown to stratify women according to their risk of developing breast cancer[6-9]. Using a large GWAS dataset from the Breast Cancer Association Consortium (BCAC), we previously developed and validated a 313-variant PRS (PRS_{313}) among women of European descent. In independent prospective studies, this PRS_{313} predicted the risk of primary invasive breast cancer with an odds ratio (OR) per standard deviation (SD) of 1.61 (95% confidence interval (95%CI)=1.57-1.65)[7]. The PRS_{313} has also been externally validated using the UK Biobank cohort.

The aim of the current study was to evaluate the association between PRS_{313} and CBC risk, using data from BCAC. Other studies have shown associations between risk of CBC and both a 67-variant PRS[10] and individual variants[11], but not yet with PRS_{313}, the most extensively validated PRS. Further, the dataset currently evaluated is larger than those previously tested. We carried out two types of analyses. We conducted a cohort study among studies of European ancestry women with follow-up data available, and performed Cox regression analyses to

estimate hazard ratios (HRs) for CBC. Potential confounding and interaction with patient characteristics, characteristics of the primary tumor, or treatment were tested. In addition, to directly compare with the OR reported for PRS_{313} and first breast cancer, we selected case-case series and performed logistic regression analyses comparing the PRS_{313} distribution in women with CBC versus those with unilateral breast cancer. These analyses were conducted separately in European and Asian women (follow-up was too limited to perform a cohort study for the Asian population). Use of PRS_{313} may lead to more accurate CBC risk prediction to support decision making for women who may or may not benefit from additional surveillance and risk-reducing treatment strategies.

Material and Methods

Study subjects

Case-case series

We selected women who were diagnosed with breast cancer and women without any diagnosis of breast cancer from the BCAC including all women of European ancestry, based on genotyping data, selecting only those studies which reported on CBC (62 studies) (Figure S1A, Table S1-S2). BCAC database version freeze 12 was used. All women diagnosed with invasive breast cancer as a first cancer were included in the analysis; the small number of tumors with unknown invasiveness were considered invasive (Table S2). In the case-case series, a CBC was defined as a breast cancer (in situ or invasive) in the contralateral breast irrespective of the time since the first breast cancer. The case-case series comprised 81,000 women with unilateral breast cancer, 3,607 women with CBC, and 62,830 women without any diagnosis of breast cancer (Figure S1A). We also compared women with unilateral breast cancer to women without any diagnosis of breast cancer to reproduce the estimate that was previously reported for first breast cancer risk[7] in our study selection.

We selected for a separate analysis women of Asian ancestry of the BCAC data comprising 12,133 women with unilateral breast cancer, 340 women with CBC, and 13,398 women without any diagnosis of breast cancer from eight studies (Figure S1B, Table S2).

European cohort

In the European cohort we used metachronous CBC as the outcome, defined as a breast cancer in the contralateral breast (in situ or invasive) diagnosed at least three months after the first breast cancer. We used a cut-off of three months to reduce the likelihood that these CBCs represent metastases rather than true second primary tumors. We selected all women diagnosed with breast cancer from the European case-case series and excluded four studies

that did not provide follow-up information on vital status (Figure S1A). We did not include Asian women since follow-up was too limited in these studies. We additionally excluded 6,207 women with no follow-up and 2,208 women who developed synchronous CBC, distant metastasis, or who died or last known to be alive within three months after the first breast cancer diagnosis. Since BCAC also included prevalent cases, we excluded 3,796 women who developed CBC or were censored before study entry. The case-case series included women diagnosed between 1947 and 2018. In the European cohort, we excluded 2,235 women who were diagnosed with their first breast cancer before 1990 or who had missing year of first diagnosis. We restricted to women diagnosed from 1990 onwards so that diagnostic procedures and treatment would be more representative of current practice. Moreover, clinico-pathological, treatment and follow-up data were more complete after 1990. In addition, we excluded 16 studies (9,783women) without information about metachronous CBC events (Figure S1A). After these exclusions, the cohort for this analysis comprised data from 42 studies, including 56,068 women with invasive breast cancer among whom 1,027 metachronous CBC occurred (Table S2).

All individuals provided written informed consent, and all studies were approved by the relevant institutional review boards. BCAC data were centrally harmonized and cleaned in communication with the study data managers and principal investigators. Data collection for individual studies is described in Table S1.

Genotyping and PRS

DNA samples from participants were genotyped using the iCOGS array[12; 13] or the OncoArray[4; 14], with genotypes for variants not on the arrays estimated by imputation[4; 13]. The PRS_{313} was calculated as a weighted sum of the minor allele dosages; the variant selection and weights are as given by Mavaddat et al.[7]. We also calculated estimates for a previously published PRS_{77}[6], and estrogen receptor (ER)-specific PRSs (ER-positive PRS_{313} and ER-negative PRS_{313})[7]. The

ER-specific PRSs were constructed by defining subtype-specific weights for the 313 variants using a hybrid approach[7]. Variants and corresponding coefficients used to construct the PRS are shown in Table S3. We standardized the PRS in our analyses by dividing it by the SD of the PRS of the controls (PRS_{77} SD=0.45; PRS_{313} SD=0.61; ER-positive PRS_{313} SD=0.65; ER-negative PRS_{313} SD=0.59) exactly as was done in the analyses of the PRS and first breast cancer risk[6; 7]. This allows a direct comparison of the magnitude of the CBC relative risk estimation to that of the first breast cancer.

For samples genotyped with both OncoArray and iCOGS array (9,071 samples), OncoArray data were used in preference as the imputation quality was generally higher. The intraclass correlation coefficient (ICC) between the PRS derived from the two platforms was 0.99 (95%CI=0.99-0.99) for the PRS_{77}, and 0.96 (95%CI=0.95-0.96) for PRS_{313} (Figure S2). Given the high correlation between the two platforms, PRS measures from both platforms were used in the analyses without adjustment.

Statistical analysis

European cohort

The primary outcome in the European cohort was the development of metachronous CBC. Cox proportional hazards models were used to estimate HRs for metachronous CBC risk by PRS, stratified by country. Since previous studies have shown that age at first breast cancer diagnosis is an important predictor of CBC[3], the analyses were performed with attained age as the time scale. Time at risk started three months after the first breast cancer diagnosis and ended at the age of CBC diagnosis, distant metastasis (where available), death, or end of follow-up, whichever came first. For patients that had a study entry more than three months after first breast cancer diagnosis, follow-up started at the age of study entry. We also performed a fixed-effect meta-analysis of country-specific effects using the STATA command

metan. We performed a fixed-effect meta-analysis over a random-effect meta-analysis since there was no evidence for heterogeneity in effect sizes between countries (I-squared=0%, Figure S3). For some analyses, only invasive CBC was used as the outcome; in these analyses we censored on in situ CBC. Separate analyses were conducted for ER-positive CBC (censored on ER-negative- and ER-unknown CBC) and ER-negative CBC (censored on ER-positive- and ER-unknown CBC).

We evaluated the linearity of the association between PRS_{313} per unit SD and CBC risk using restricted cubic splines with three knots. There was no evidence for violation of the linearity assumption. Therefore, in the main analysis, the PRS_{313} was treated as a continuous covariate, and estimated the HR per unit SD of the $PRS_{313.}$ Violation of the proportional hazard assumption was assessed by inspection of the Schoenfeld residuals[15]. As a second analysis, we used the per SD log HR of the PRS_{313} to calculate the predicted HR at different percentiles of the PRS_{313}, compared to the 50^{th} percentile. Third, the PRS_{313} was categorized into percentile groups (0^{th} to 10^{th}, 10^{th} to 20^{th}, 20^{th} to 40^{th}, 40^{th} to 60^{th}, 60^{th} to 80^{th}, 80^{th} to 90^{th}, 90^{th} to 100^{th}) to illustrate the differences between PRS_{313} subgroups, with the middle quintile (40^{th} to 60^{th}) as the reference.

We also performed multivariable Cox regression analyses to determine whether the log HR of CBC risk by PRS changed when adjusting for year of first breast cancer diagnosis, family history of breast cancer in a first degree relative, and several clinical characteristics of the first breast cancer such as nodal status, tumor size, morphology, ER-, progesterone receptor (PR)- and human epidermal growth factor receptor 2 (HER2)-status, (neo)adjuvant chemotherapy, adjuvant endocrine therapy, and radiotherapy. These analyses were performed in all patients, a complete case set (excluding patients with unknown values for the covariates), and in a set excluding studies oversampling cases with family history. Potential effect modification of the

PRS_{313} effect by the same variables was evaluated by fitting interaction terms in different models using complete case sets, including the standardized PRS_{313}, modifier, and interaction.

The discriminative ability of different models; ([model 1] PRS_{313} alone, [model 2] other risk factors (the adjustment variables from the multivariable Cox regression analyses), [model 3] PRS_{313} + other risk factors) was calculated using Harrell's C-index[16]. Since no standard performance measures are currently available to account for left-truncated follow-up time (*i.e*., to start analyses at age at study entry), we used time since first breast cancer as the time scale to calculate the C-index.

Absolute risks

Absolute risks of developing CBC at PRS_{313} percentiles were calculated using the estimated log HRs per SD from the breast cancer cohort (BCAC) under the log-linear model, assuming the PRS is normally distributed. The PRS_{313}- and age-specific incidences were constrained to the age-specific CBC incidences from women diagnosed with a first invasive breast cancer in the period 2003-2010 from the Netherlands Cancer Registry (NCR)[1]. The procedure for constraining the incidences has been previously described[17]. The age-specific CBC incidences were calculated overall and for age-specific groups, censoring on death and distant metastasis. We used data from the NCR since this registry has complete coverage of all newly diagnosed cancers in the Netherlands. The NCR cohort included all females aged ≥18 years and follow-up for second cancers was complete until February 1, 2016[1]. We then applied the competing risk of dying on the absolute CBC risks. The absolute CBC risk (AR_g) by age t in PRS_{313} category g, taking into account the competing risk of dying was calculated by:

$$AR_g(t) = \sum_{u=0}^{t-1} \mu_g(u) S_g(u) S_m(u)$$

Where $\mu_g\ (t)$ is the CBC incidence associated with PRS_{313} category g, $S_g\ (t)$ the probability of being free of CBC to age t, and $S_m\ (t)$ the probability of surviving to age t.

Case-case series

For the case-case series (European and Asian), logistic regression models were used to estimate the ORs for CBC risk (comparing with unilateral breast cancer) and for unilateral breast cancer risk (comparing with women without any diagnosis of breast cancer) associated with PRS_{313}. All analyses were adjusted for age and country (Table S1). For all unilateral- and contralateral breast cancer patients we used age at first breast cancer diagnosis, and for women without any diagnosis of breast cancer we used age at baseline questionnaire.

For direct comparison with the estimate reported for PRS_{313} and first breast cancer, we also performed logistic regression analyses in the same BCAC study participants included in the validation of the association between PRS_{313} and first breast cancer risk[7]. This validation set comprised a subsample from 24 studies and included 3,781 women with unilateral breast cancer, 94 women with CBC, and 3,753 women without any diagnosis of breast cancer (Table S2). For this analysis, we adjusted for 10 principal components, in line with Mavaddat et al.[7].

For European women who had follow-up time available more than three months after the first breast cancer diagnosis, a sensitivity analysis was performed for metachronous CBC (1,702 CBCs). We also did a separate analysis for invasive CBC (N=3,246), by excluding CBC in situ.

All P-values are two sided; tests with P<.05 are referred to as statistically significant. Analyses were performed using STATA, version 13.1 (StataCorp) and R version 3.3.2.

Results

European (cohort) Cox regression analyses

The European cohort included 56,068 women diagnosed with first invasive breast cancer with 1,027 metachronous CBC events. Median follow-up was 8.4 years. Patient, tumor, and treatment characteristics are summarized in Table S4.

The associations between the different PRSs and CBC risk are shown in Table 1. The HR for CBC per SD of PRS_{313} was 1.25 (95%CI=1.18-1.33). For comparison, the HR per SD for PRS_{77} was 1.21 (95%CI=1.14-1.29). Women within the 0^{th} to 10^{th} and the 90^{th} to 100^{th} percentile of the PRS_{313} had 0.59-fold (95%CI=0.45-0.78) and 1.38-fold (95%CI=1.13-1.69) risks of CBC, respectively, compared with women within the 40^{th} to 60^{th} percentile (Figure 1, Table S5). The predicted HRs of CBC for women at the 10^{th} and 90^{th} percentile of the PRS_{313} were 0.75 and 1.33, respectively, compared to the 50^{th} percentile (Figure 1). Since we observed evidence of departure from the proportional hazards assumption (P=0.02)[15], we also calculated HRs stratified for follow-up duration (<five and ≥five years). The HR by SD of the PRS_{313} was 1.21 (95%CI=1.10-1.32) for CBC diagnosed ≤five years after first breast cancer diagnosis (CBC N=428), and 1.28 (95%CI=1.18-1.38) for CBC diagnosed >five years after first diagnosis (CBC N=599).

The HR per SD of PRS_{313} for ER-positive invasive CBC was 1.38 (95%CI=1.23-1.55), compared to a HR per SD of the ER-positive PRS_{313} of 1.37 (95%CI=1.22-1.54) (Table 1). For ER-negative invasive CBC, the HR per SD was 0.92 (95%CI=0.75-1.12) for PRS_{313} and 1.06 (95%CI=0.86-1.30) for the ER-negative PRS_{313}.

Sensitivity analysis using the overall PRS_{313} showed a HR per SD of 1.24 (95%CI=1.16-1.32) for invasive CBC risk. When we used time since first breast cancer as the time scale, we found

similar results (HR per SD=1.25, 95%CI=1.18-1.33). Meta-analysis of country-specific effects showed a HR per SD of 1.25 (95%CI=1.18-1.33) for CBC risk by PRS_{313} (Figure S3).

The association between the PRS_{313} and CBC risk did not change when adjusting for patient, tumor, and treatment characteristics, nor when excluding studies oversampling cases with a family history (Table S6). When considering potential modifiers of the effect of the PRS_{313} on CBC risk (Table 2), we found that the HR was the lowest in women aged <40 years at first breast cancer diagnosis (HR per SD=1.13; 95%CI=0.98-1.31), and tended to increase with age, although these effects were not statistically significant ($P_{heterogeneity}$=.26; P_{trend}=.05). We found no indication for effect modification by family history ($P_{heterogeneity}$=.63), morphology ($P_{heterogeneity}$=.14), ER-status ($P_{heterogeneity}$=.13), PR-status (P=.26), HER2-status ($P_{heterogeneity}$=.42), chemotherapy ($P_{heterogeneity}$=.60), endocrine therapy ($P_{heterogeneity}$=.79), or radiotherapy ($P_{heterogeneity}$ =.40) (Table 2).

The C-index was 0.563 (95%CI=0.547-0.586) for the model only including PRS_{313}, 0.605 (95%CI=0.591-0.629) for the model only including other risk factors, and 0.623 (95%CI=0.608-0.645) for the complete model (Table 3).

Absolute risks

Based on the HR estimates for PRS_{313}, the predicted CBC risk by age 80 years was 12.4% at the 10^{th} percentile of the PRS_{313}, compared with 20.5% at the 90^{th} percentile of the PRS_{313} (Figure 2), accounting for death as competing risk. When death was not taken into account as competing risk, the corresponding predicted risks by age 80 were 17.0% at the 10% percentile and 27.9% at the 90^{th} percentile of the PRS_{313} (Figure S4). Table 4 shows the five- and 10-year cumulative CBC risks by PRS_{313} for different age groups, accounting for death as competing risk (Table S7 shows results without competing risks).

European and Asian (case-case series) logistic regression analyses

Figure 3 shows the distribution of the PRS_{313} per SD in the European case-case series. Median PRS_{313} was -0.4 (interquartile range [IQR]=1.35) for control women without any diagnosis of breast cancer (N=81,000), 0.2 (IQR=1.36) for women with unilateral breast cancer (N=62,830), and 0.5 (IQR=1.40) for women with CBC (N=3,607). The OR for unilateral breast cancer per SD of the PRS_{313}, compared to control women, was 1.82 (95%CI=1.80-1.84) (Table S8). The OR for CBC per SD of PRS_{313}, compared to unilateral breast cancer, was 1.30 (95%CI=1.26-1.35) .

In sensitivity analyses, the OR per SD of PRS_{313} was 1.27 (95%CI=1.21-1.33) for metachronous CBC and the OR per SD was 1.29 (95%CI=1.24-1.33) for invasive CBC, compared to unilateral breast cancer. When analyses were restricted to the validation set of Mavaddat et al[7], the OR for unilateral breast cancer per SD of the PRS_{313} was 1.67 (95%CI=1.59-1.76) compared to control women, and the OR for CBC per SD of PRS_{313} was 1.39 (95%CI=1.13-1.70) compared to unilateral breast cancer (Table S8).

For women of Asian descent, the OR for unilateral breast cancer per SD of the PRS_{313} was 1.56 (95%CI=1.52-1.60) compared to control women, and the OR for CBC per SD of PRS_{313} was 1.15 (95%CI=1.02-1.29) compared to women with unilateral breast cancer (Table S8).

Discussion

Previous studies have shown that a PRS, summarizing the effects of common germline variants, can be used to stratify women with respect to their risk to develop a primary breast cancer[6-9]. In this study, we observed a clear association between the PRS_{313} and CBC risk in women of both European and Asian ancestry. The association was observed in both the case-case series and the European cohort. The HRs per SD of CBC for women at the 10^{th} and 90^{th} percentile of the continuous predicted PRS_{313} were 0.75 and 1.33, respectively, compared to the 50^{th} percentile. This translates to absolute risks at the 10^{th} and the 90^{th} percentile of the PRS_{313} of 12.4% and 20.5%, respectively, by age 80 years. We estimated a C-index for the PRS_{313}, summarizing its discriminatory ability, of 0.563 in the European cohort.

One previous study has investigated the effect of a PRS, including 67 variants, and CBC risk[10]. This study found a risk ratio of 1.75 (95%CI=1.41-2.18) for women in the upper quartile of the PRS compared with women in the lowest quartile. To facilitate comparison, we performed a similar analysis in our case-case series, showing an OR of 1.98 (95%CI=1.79-2.18), adjusted for country and age at first diagnosis, for women in the upper quartile of the PRS_{313}. This indicates the PRS_{313} improves stratification relative to PRSs including fewer variants. Moreover, in our European cohort, the C-index for the PRS alone improved from 0.547 (95%CI=0.536-0.575) for the previously reported PRS_{77}[6] to 0.563 (95%CI=0.547-0.586) for the PRS_{313}.

We found no evidence that the association between the PRS_{313} and CBC risk was confounded by family history, adjuvant therapy, morphology, age, or tumor receptor status of the first breast cancer, nor that there was effect modification by those factors. The absence of notable effect modification is in line with the abovementioned study of a 67-variant PRS and CBC risk; no heterogeneity in association was found by age, family history, morphology, ER-status, and adjuvant treatment[10].

To provide an external validation of our findings, we examined data from UK Biobank, which includes many women diagnosed with breast cancer with data available on the PRS_{313} (Supplemental Note). Unfortunately, UK Biobank has no information available on the laterality of the tumor, and it is, therefore, not possible to distinguish between contralateral and ipsilateral breast cancers. We therefore performed analyses using any second breast cancer as the endpoint. This secondary analysis did confirm the association between the PRS_{313} and second breast cancer risk (HR per SD=1.13, 95%CI=1.01-1.27), but with a lower estimate than in our European cohort. The lower estimate may be explained by the inclusion of the ipsilateral breast cancers, which may be more likely to be recurrences than new primary breast cancers compared to CBCs. Indeed, when we used ipsilateral breast cancer as the outcome in our European cohort, we found no association with the PRS_{313} (HR=1.02, 95%CI=0.90-1.15).

The association between the PRS_{313} and CBC risk (OR per SD=1.30; 95%CI=1.26-1.35) in the BCAC database was weaker (expressed in terms of an OR) than was found for first breast cancer among independent prospective studies (OR per SD=1.61; 95%CI=1.57-1.65). Under a simple polygenic model, the relative risk would be expected to be similar for the second breast cancer. The attenuated estimate for CBC might however be explained by several factors. Some attenuation of the estimate might have been due to dilution in the end-point definition, *i.e.*, if some of the CBCs were metastases. Previous studies investigating the clonal relatedness of first breast cancers and CBCs using tumor sequencing have shown that 6-12% of CBCs represent metastases[18; 19]. This hypothesis would be consistent with our finding of a slightly stronger association between the PRS_{313} and late CBCs, diagnosed >five years after the first breast cancer, than for early CBCs, diagnosed ≤five years after the first cancer, since the latter are more likely to be metastases. In addition, 3-5% of the breast cancer patients will have a mutation in the *BRCA1 or BRCA2 gene*[20; 21]*,* who have high CBC risks. It has been shown that

the relative risk associated with PRS is lower (for the first breast cancer) for women with a *BRCA1* and *BRCA2* mutation than in the general population[22], diluting the overall relative risk for CBC. More generally, it is possible that the CBC association may be attenuated due to the effect of other, unmeasured, genetic or other risk factors. If the risks are high, cases with higher PRS_{313} will have, on average, lower values of other risk factors, due to elimination of the highest risk individuals, again attenuating the CBC association. Finally, given the limited information on family history in our dataset, the estimate could have been biased due to a family history effect not detected in our data.

There was some suggestion that the relative risk associated with PRS_{313} decreased with younger age, (P_{trend}=.05), and, specifically, was lower for women aged <40 years (HR per SD=1.13; 95%CI=0.98-1.31). Interestingly, Mavaddat et al[7] also found a lower relative risk below age 40 for first breast cancer. This effect may reflect the different characteristics of breast cancers at young ages, both in terms of germline susceptibility and pathology[23; 24]. For example, the proportion of ER-negative breast cancers is higher at young ages, and the PRS is less predictive for ER-negative disease[6; 7; 24].

In the logistic regression analyses in Asian women, the association between the PRS_{313} and CBC risk was slightly weaker than in European women. This finding is consistent with a recent analysis investigating the association between a 287-variant PRS and first breast cancer risk in the Asian population[25], which showed an attenuated OR in Asian women (OR=1.52, 95%CI=1.49-1.56) compared to European women (OR=1.61, 95%CI=1.57-1.66). The lower estimate for Asian women might reflect the fact the PRS_{313} was developed in European populations, and the different LD structure in Asians may attenuate the association since the variants in the PRS are likely to be surrogates for the causal variants. Other explanations for the attenuated estimate may be the slightly younger age at first breast cancer diagnosis and the

higher proportion ER-negative CBCs in Asian women compared to European women in our study. Finally, the imputation quality for variants was somewhat lower, on average, for the Asian than for the European dataset, with three variants on OncoArray and four variants on ICOGs with an imputation quality score<0.3 (Table S3). Nevertheless, we included those variants in the PRS for both European and Asian women, to keep the PRS comparable between ethnicities and studies. Future studies including larger numbers of Asian women, and women of other ethnicities, are needed to generate population-specific PRSs and to validate our findings in these groups.

A major strength of this study is the very large sample size in the BCAC dataset, including genotype information for ~150,000 women and a large number of CBC events. A limitation of this study is missing data on the patient, tumor, and treatment characteristics, which reduces the power of the multivariable Cox regression analyses and interaction analyses. In addition, registration of CBC was not complete; the 10-year cumulative CBC incidence was 2.2% in the BCAC dataset, compared to 3.8% using complete data from the Netherlands Cancer Registry[1]. For this reason, we estimated relative risk estimates using the BCAC data and applied these to external registry data to obtain absolute risk estimates. The underreporting of CBC should not bias our HR estimates, given that the event rate is low and reporting of CBC is unlikely to be related to the PRS_{313}. Moreover, we reran the cohort analysis in the subset of countries with a 10-year cumulative CBC incidence ≥3.0% in the BCAC dataset, and the estimates were very similar to the main analyses (HR per SD=1.23, 95%CI=1.14-1.33) (Figure S3).

In conclusion, the PRS_{313} is predictive for the development of CBC. We found no evidence for confounding or effect modification by other previously established CBC risk factors. The PRS_{313} is therefore likely to be an independent risk factor for CBC. Since the predictive ability of the PRS on its own is modest, it should be combined with other breast cancer risk factors to provide

more useful CBC risk prediction models. More accurate risk prediction will help identify women at high CBC risk who will benefit from additional surveillance and/or risk reducing mastectomy, and equally important, to identify those women at low risk in order to avoid unnecessary surgeries.

Supplemental Data

Supplemental data include four figures, eight tables, Supplemental Note and acknowledgements.

Data and Code Availability

Data used in this manuscript may be requested through the original providers. Data of the Breast Cancer Association Consortium may be requested for non-profit research through an application procedure with the Breast Cancer Association Consortium; more information: http://bcac.ccge.medschl.cam.ac.uk/bcacdata/. Data of the UK Biobank needs to be requested through UK Biobank; more information: https://www.ukbiobank.ac.uk/researchers/

Acknowledgements

Acknowledgements and funding are included in the Supplement.

Declaration of Interests

Dr. Beckmann conducts research funded by Amgen, Novartis and Pfizer, outside the submitted work. Dr. Fasching conducts research funded by Amgen, Novartis and Pfizer, outside the submitted work. He received honoraria from Roche, Novartis and Pfizer. Dr. Nevanlinna received honorarium from Astra Zeneca outside the submitted work.

References

1. Kramer, I., Schaapveld, M., Oldenburg, H.S.A., Sonke, G.S., McCool, D., van Leeuwen, F.E., Van de Vijver, K.K., Russell, N.S., Linn, S.C., Siesling, S., et al. (2019). The influence of adjuvant systemic regimens on contralateral breast cancer risk and receptor subtype. J Natl Cancer Inst.

2. Xiong, Z., Yang, L., Deng, G., Huang, X., Li, X., Xie, X., Wang, J., Shuang, Z., and Wang, X. (2018). Patterns of Occurrence and Outcomes of Contralateral Breast Cancer: Analysis of SEER Data. Journal of clinical medicine 7.

3. Akdeniz, D., Schmidt, M.K., Seynaeve, C.M., McCool, D., Giardiello, D., van den Broek, A.J., Hauptmann, M., Steyerberg, E.W., and Hooning, M.J. (2019). Risk factors for metachronous contralateral breast cancer: A systematic review and meta-analysis. Breast (Edinburgh, Scotland) 44, 1-14.

4. Michailidou, K., Lindström, S., Dennis, J., Beesley, J., Hui, S., Kar, S., Lemaçon, A., Soucy, P., Glubb, D., Rostamianfar, A., et al. (2017). Association analysis identifies 65 new breast cancer risk loci. Nature 551, 92.

5. Milne, R.L., Kuchenbaecker, K.B., Michailidou, K., Beesley, J., Kar, S., Lindstrom, S., Hui, S., Lemacon, A., Soucy, P., Dennis, J., et al. (2017). Identification of ten variants associated with risk of estrogen-receptor-negative breast cancer. Nature genetics 49, 1767-1778.

6. Mavaddat, N., Pharoah, P.D., Michailidou, K., Tyrer, J., Brook, M.N., Bolla, M.K., Wang, Q., Dennis, J., Dunning, A.M., Shah, M., et al. (2015). Prediction of breast cancer risk based on profiling with common genetic variants. J Natl Cancer Inst 107.

7. Mavaddat, N., Michailidou, K., Dennis, J., Lush, M., Fachal, L., Lee, A., Tyrer, J.P., Chen, T.H., Wang, Q., Bolla, M.K., et al. (2019). Polygenic Risk Scores for Prediction of Breast Cancer and Breast Cancer Subtypes. American journal of human genetics 104, 21-34.

8. Brentnall, A.R., van Veen, E.M., Harkness, E.F., Rafiq, S., Byers, H., Astley, S.M., Sampson, S., Howell, A., Newman, W.G., Cuzick, J., et al. (2019). A case-control evaluation of 143 single

nucleotide polymorphisms for breast cancer risk stratification with classical factors and mammographic density. International journal of cancer.

9. Shieh, Y., Hu, D., Ma, L., Huntsman, S., Gard, C.C., Leung, J.W., Tice, J.A., Vachon, C.M., Cummings, S.R., Kerlikowske, K., et al. (2016). Breast cancer risk prediction using a clinical risk model and polygenic risk score. Breast cancer research and treatment 159, 513-525.

10. Robson, M.E., Reiner, A.S., Brooks, J.D., Concannon, P.J., John, E.M., Mellemkjaer, L., Bernstein, L., Malone, K.E., Knight, J.A., Lynch, C.F., et al. (2017). Association of Common Genetic Variants With Contralateral Breast Cancer Risk in the WECARE Study. JNCI: Journal of the National Cancer Institute 109, djx051-djx051.

11. Teraoka, S.N., Bernstein, J.L., Reiner, A.S., Haile, R.W., Bernstein, L., Lynch, C.F., Malone, K.E., Stovall, M., Capanu, M., Liang, X., et al. (2011). Single nucleotide polymorphisms associated with risk for contralateral breast cancer in the Women's Environment, Cancer, and Radiation Epidemiology (WECARE) Study. Breast cancer research : BCR 13, R114.

12. Michailidou, K., Beesley, J., Lindstrom, S., Canisius, S., Dennis, J., Lush, M.J., Maranian, M.J., Bolla, M.K., Wang, Q., Shah, M., et al. (2015). Genome-wide association analysis of more than 120,000 individuals identifies 15 new susceptibility loci for breast cancer. Nature genetics 47, 373-380.

13. Michailidou, K., Hall, P., Gonzalez-Neira, A., Ghoussaini, M., Dennis, J., Milne, R.L., Schmidt, M.K., Chang-Claude, J., Bojesen, S.E., Bolla, M.K., et al. (2013). Large-scale genotyping identifies 41 new loci associated with breast cancer risk. Nature genetics 45, 353-361, 361e351-352.

14. Amos, C.I., Dennis, J., Wang, Z., Byun, J., Schumacher, F.R., Gayther, S.A., Casey, G., Hunter, D.J., Sellers, T.A., Gruber, S.B., et al. (2017). The OncoArray Consortium: A Network for Understanding the Genetic Architecture of Common Cancers. Cancer epidemiology, biomarkers & prevention : a publication of the American Association for Cancer Research, cosponsored by the American Society of Preventive Oncology 26, 126-135.

15. Schoenfeld, D.A. (1983). Sample-size formula for the proportional-hazards regression model. Biometrics 39, 499-503.

16. Harrell, F.E., Jr., Califf, R.M., Pryor, D.B., Lee, K.L., and Rosati, R.A. (1982). Evaluating the yield of medical tests. Jama 247, 2543-2546.

17. Antoniou, A.C., Beesley, J., McGuffog, L., Sinilnikova, O.M., Healey, S., Neuhausen, S.L., Ding, Y.C., Rebbeck, T.R., Weitzel, J.N., Lynch, H.T., et al. (2010). Common breast cancer susceptibility alleles and the risk of breast cancer for BRCA1 and BRCA2 mutation carriers: implications for risk prediction. Cancer research 70, 9742-9754.

18. Klevebring, D., Lindberg, J., Rockberg, J., Hilliges, C., Hall, P., Sandberg, M., and Czene, K. (2015). Exome sequencing of contralateral breast cancer identifies metastatic disease. Breast cancer research and treatment 151, 319-324.

19. Begg, C.B., Ostrovnaya, I., Geyer, F.C., Papanastasiou, A.D., Ng, C.K.Y., Sakr, R.A., Bernstein, J.L., Burke, K.A., King, T.A., Piscuoglio, S., et al. (2018). Contralateral breast cancers: Independent cancers or metastases? International journal of cancer 142, 347-356.

20. Thompson, D., and Easton, D. (2004). The genetic epidemiology of breast cancer genes. Journal of mammary gland biology and neoplasia 9, 221-236.

21. van den Broek, A.J., van 't Veer, L.J., Hooning, M.J., Cornelissen, S., Broeks, A., Rutgers, E.J., Smit, V.T., Cornelisse, C.J., van Beek, M., Janssen-Heijnen, M.L., et al. (2016). Impact of Age at Primary Breast Cancer on Contralateral Breast Cancer Risk in BRCA1/2 Mutation Carriers. Journal of clinical oncology : official journal of the American Society of Clinical Oncology 34, 409-418.

22. Kuchenbaecker, K.B., McGuffog, L., Barrowdale, D., Lee, A., Soucy, P., Dennis, J., Domchek, S.M., Robson, M., Spurdle, A.B., Ramus, S.J., et al. (2017). Evaluation of Polygenic Risk Scores for Breast and Ovarian Cancer Risk Prediction in BRCA1 and BRCA2 Mutation Carriers. J Natl Cancer Inst 109.

23. Azim, H.A., Jr., Michiels, S., Bedard, P.L., Singhal, S.K., Criscitiello, C., Ignatiadis, M., Haibe-Kains, B., Piccart, M.J., Sotiriou, C., and Loi, S. (2012). Elucidating prognosis and biology of breast cancer arising in young women using gene expression profiling. Clinical cancer research : an official journal of the American Association for Cancer Research 18, 1341-1351.
24. Anders, C.K., Hsu, D.S., Broadwater, G., Acharya, C.R., Foekens, J.A., Zhang, Y., Wang, Y., Marcom, P.K., Marks, J.R., Febbo, P.G., et al. (2008). Young age at diagnosis correlates with worse prognosis and defines a subset of breast cancers with shared patterns of gene expression. Journal of clinical oncology : official journal of the American Society of Clinical Oncology 26, 3324-3330.
25. Ho, W.K., Tan, M.M., Mavaddat, N., Tai, M.C., Mariapun, S., Li, J., Ho, P.J., Dennis, J., Tyrer, J.P., Bolla, M.K., et al. (2020). European polygenic risk score for prediction of breast cancer shows similar performance in Asian women. Nature communications 11, 3833.

Figure 1. Estimates for contralateral breast cancer risk by percentile categories of the 313-variant PRS (PRS_{313})

The figure shows the hazard ratios per SD and 95% confidence intervals for percentiles of the PRS_{313} relative to the middle quintile (underlying table can be found in Table S5). The solid line denotes the estimates for contralateral breast cancer risk with the PRS_{313} fitted as a continuous covariate. Coefficients to construct the PRS_{313} are shown in Table S3. The PRS_{313} was standardized by SD=0.61, in line with Mavaddat et al.[7]. The analyses were performed with attained age as time scale. PRS = polygenic risk score, SD = standard deviation

Figure 2. Predicted contralateral breast cancer risk by percentile of the 313-variant PRS (PRS_{313}) with death as competing risk

Coefficients to construct the PRS_{313} are shown in Table S3. The PRS_{313} was standardized by SD=0.61, in line with Mavaddat et al.[7] The CBC incidences were calculated based on incidence data from the Netherlands Cancer Registry[1] and relative risks estimated as described in the Material and Methods. PRS = polygenic risk score, CBC = contralateral breast cancer

Figure 3. Distribution of the 313-variant PRS (PRS_{313}) in 62,830 control women without any diagnosis of breast cancer, 81,000 women with unilateral breast cancer, and 3,607 women with contralateral breast cancer

Coefficients to construct the PRS_{313} are shown in Table S3. The PRS_{313} was standardized by SD=0.61, in line with Mavaddat et al.[7]. PRS = polygenic risk score, BC = breast cancer, CBC = contralateral breast cancer, SD = standard deviation

Table 1. Association between PRSs and contralateral breast cancer risk in the European cohort (N=56,068)

Polygenic risk score (PRS)	No. of CBC	HR per unit SD[a]	95%CI	P-value
PRS_{77} [b]				
All CBC	1,027	1.21	1.14-1.29	<.001
Invasive CBC	923	1.21	1.13-1.29	<.001
PRS_{313} [b]				
All CBC	1,027	1.25	1.18-1.33	<.001
Invasive CBC	923	1.24	1.16-1.32	<.001
ER-positive invasive CBC[d]	275	1.38	1.23-1.55	<.001
ER-negative invasive CBC[d]	97	0.92	0.75-1.12	.39
ER-positive PRS_{313} [b,c]				
All CBC	1,027	1.23	1.16-1.31	<.001
Invasive CBC	923	1.22	1.15-1.30	<.001
ER-positive invasive CBC[d]	275	1.37	1.22-1.54	<.001
ER-negative PRS_{313} [b,c]				
All CBC	1,027	1.25	1.17-1.33	<.001
Invasive CBC	923	1.24	1.16-1.33	<.001
ER-negative invasive CBC[d]	97	1.06	0.86-1.30	.58

Abbreviations: PRS = polygenic risk score, No. = number, CBC = contralateral breast cancer, HR = hazard ratio, CI = confidence interval, ER = estrogen receptor, SD = standard deviation

[a] All analyses were performed with attained age as time scale

[b] Coefficients to construct the PRSs are shown in Table S3. All PRSs were standardized by the same SD as was used by Mavaddat et al.[7]. The SD was 0.45 for overall breast cancer PRS_{77}, 0.61 for overall breast cancer PRS_{313}, 0.65 for ER-positive PRS_{313}, and 0.59 for ER-negative PRS_{313}

[c] ER-specific PRSs were constructed using a hybrid method, as described by Mavaddat et al.[7]

[d] Patients with ER-unknown CBC (N=551) were censored in these analyses

Table 2. Association between the 313-variant PRS (PRS_{313}) and contralateral breast cancer risk for subgroups

Subgroups	No. of patients	No. of CBC	HR per unit SD[a,b]	95%CI	P-value	$P_{heterogeneity}$[c,d]	P_{trend}[c,e]
All patients	56,068	1,027	1.25	1.18-1.33	<.001	-	-
Age at first breast cancer diagnosis (years)						.26	.05
<40	5,877	171	1.13	0.98-1.31	.09		
40-49	11,928	265	1.25	1.11-1.41	<.001		
50-59	16,882	320	1.22	1.09-1.36	<.001		
60+	21,381	271	1.36	1.21-1.52	<.001		
Family history (first degree relative)						.63	-
no	33,623	618	1.26	1.16-1.36	<.001		
yes	10,369	302	1.22	1.09-1.36	<.001		
Morphology						.14	-
ductal	37,324	621	1.21	1.12-1.31	<.001		
lobular	5,878	118	1.32	1.10-1.59	.002		
mixed (ductal and lobular)	2,174	46	1.52	1.15-2.02	.004		
other	3,344	70	1.20	0.96-1.50	.11		
ER-status						.13	-
negative	9,527	194	1.13	0.98-1.30	.08		
positive	38,090	670	1.28	1.19-1.38	<.001		
PR-status						.26	-
negative	13,098	244	1.16	1.03-1.32	.02		
positive	27,044	554	1.27	1.17-1.38	<.001		
HER2-status						.42	-
negative	23,787	352	1.29	1.17-1.44	<.001		
positive	4,969	60	1.45	1.13-1.85	.004		
(Neo)adjuvant chemotherapy						.60	-
no	18,110	361	1.28	1.16-1.42	<.001		
yes	18,559	363	1.24	1.12-1.37	<.001		
(Neo)adjuvant endocrine therapy						.79	-
no	10,781	242	1.28	1.13-1.44	<.001		
yes	27,322	460	1.30	1.19-1.43	<.001		
Radiotherapy						.40	-
no	11,023	188	1.33	1.15-1.53	<.001		
yes	29,142	617	1.24	1.15-1.34	<.001		

Abbreviations: PRS = polygenic risk score, No. = number, CBC = contralateral breast cancer, HR = hazard ratio, CI = confidence interval, ER = estrogen receptor, PR = progesterone receptor, HER2 = human epidermal growth factor receptor 2

[a] HR for CBC risk by unit SD of PRS_{313}. All analyses were performed with attained age as time scale

[b] Coefficients to construct the PRS_{313} are shown in Table S3. The PRS_{313} was standardized by standard deviation=0.61, in line with Mavaddat et al.[7]

[c] The interaction between the PRS_{313} and each subgroup was tested in different models including the standardized PRS_{313}, modifier, and interaction. Patients with unknown values were excluded from these analyses. Since attained age was used as time scale in all models, the model with age at first breast cancer only included the PRS_{313} and interaction

[d] P for interaction based on test for heterogeneity across categories

[e] P for interaction based on a trend test with age as continuous variable

Table 3. Discriminatory ability (C-index) of the 313-variant PRS (PRS_{313}) and other risk factors for contralateral breast cancer risk in the European cohort

	C-index (95%CI)[a,b]
Model 1 PRS_{313}[c] alone	0.563 (0.547-0.586)
Model 2 Other risk factors[d]	0.605 (0.591-0.629)
Model 3 PRS_{313}[c] + other risk factors[d]	0.623 (0.608-0.645)

Abbreviations: PRS = polygenic risk score, CI = confidence interval

[a] The Harrell's C-index was obtained by the STATA stcox postestimation command 'estat concordance', using time since first breast cancer on the time scale without taking delayed entry (prevalent cases) into account. We did not consider delayed-entry since no standard performance measures are currently available in the statistical literature to account for left-truncated follow-up time. The median of delayed entry was 0.4 years (standard deviation=2.7) in our study

[b] The 95% CIs were obtained by use of the 'somersd' package in STATA

[c] Coefficients to construct the PRS_{313} are shown in Table S3. The PRS_{313} was standardized by SD=0.61, in line with Mavaddat et al.[7]

[d] Including age at first diagnosis, year of first diagnosis, family history for breast cancer in a first degree relative, and clinical characteristics of the first breast cancer (nodal status, tumor size, differentiation grade, morphology, estrogen receptor status, human epidermal growth factor receptor 2 status, chemotherapy, endocrine therapy, radiotherapy)

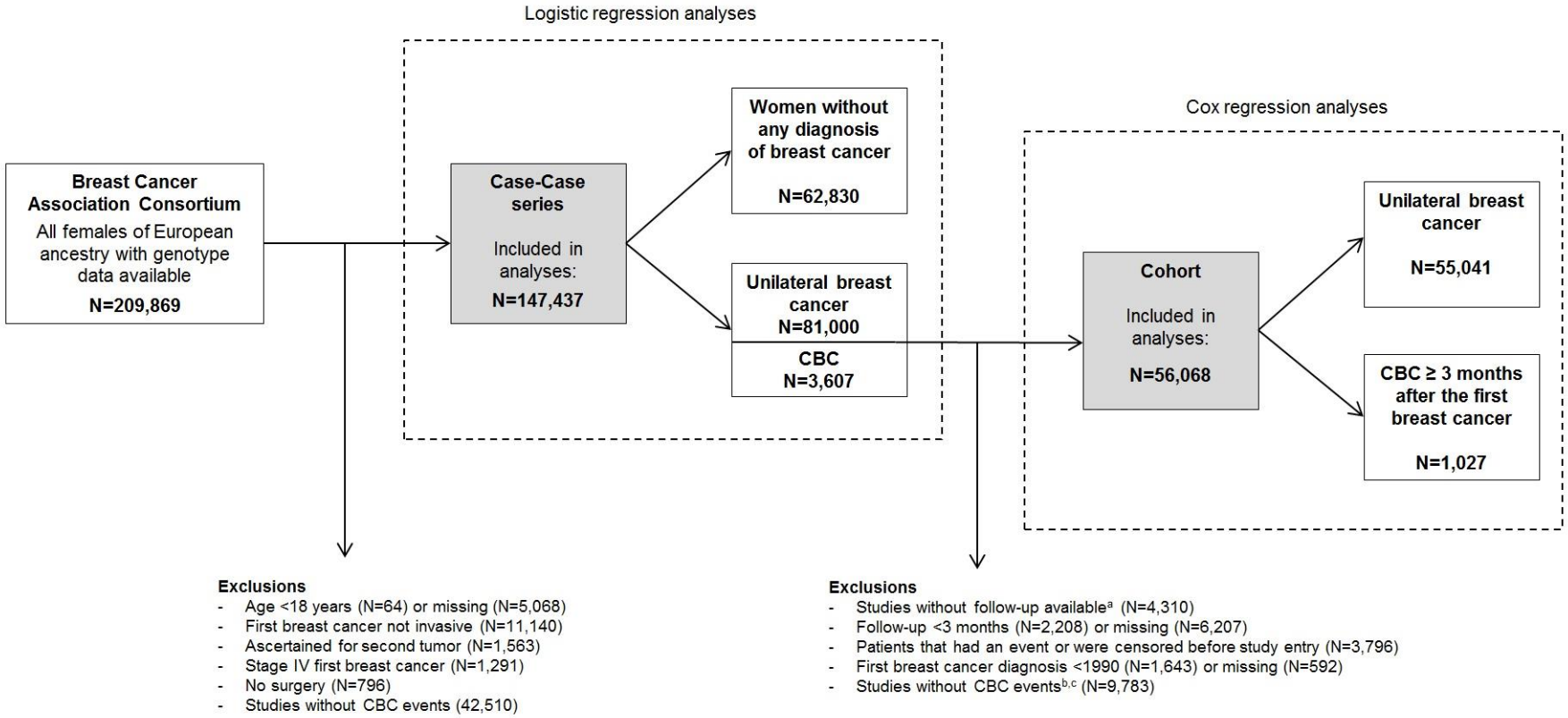
Table 4. Five- and ten-year cumulative risks of contralateral breast cancer by the 313-variant PRS (PRS_{313}) for different age groups with death as competing risk

	5-year cumulative CBC risks (%) range by age					10-year cumulative CBC risks (%) range by age				
Age at first breast cancer diagnosis (years)	**5^{th} percentile PRS_{313}**	**10^{th} percentile PRS_{313}**	**50^{th} percentile PRS_{313}**	**90^{th} percentile PRS_{313}**	**95^{th} percentile PRS_{313}**	**5^{th} percentile PRS_{313}**	**10^{th} percentile PRS_{313}**	**50^{th} percentile PRS_{313}**	**90^{th} percentile PRS_{313}**	**95^{th} percentile PRS_{313}**
30-34	1.9-3.1	2.1-3.4	2.7-4.5	3.6-5.9	4.0-6.5	3.1-4.1	3.4-4.5	4.5-5.9	5.9-7.7	6.5-8.5
35-39	0.8-2.1	0.9-2.3	1.2-3.0	1.5-3.9	1.7-4.3	2.1-3.5	2.3-3.8	3.0-5.0	3.9-6.6	4.3-7.2
40-44	1.5-2.8	1.7-3.1	2.2-4.1	2.9-5.3	3.2-5.9	2.8-4.6	3.1-5.0	4.1-6.6	5.3-8.6	5.9-9.4
45-49	1.4-2.5	1.5-2.7	2.0-3.6	2.6-4.7	2.9-5.2	2.5-3.9	2.7-4.3	3.6-5.6	4.7-7.4	5.2-8.1
50-54	1.4-2.8	1.5-3.0	1.9-4.0	2.6-5.2	2.8-5.8	2.8-4.5	3.0-4.9	4.0-6.4	5.2-8.4	5.8-9.3
55-59	1.6-3.1	1.8-3.4	2.3-4.5	3.1-5.9	3.4-6.5	3.1-4.8	3.4-5.2	4.5-6.9	5.9-9.0	6.5-9.9
60-64	1.7-3.3	1.9-3.6	2.5-4.7	3.3-6.2	3.6-6.8	3.3-5.0	3.6-5.4	4.7-7.1	6.2-9.3	6.8-10.2
65-70	1.5-3.2	1.6-3.5	2.1-4.6	2.8-6.1	3.1-6.7	3.2-4.1	3.5-4.5	4.6-5.9	6.1-7.7	6.7-8.5

Abbreviations: PRS = polygenic risk score, CBC = contralateral breast cancer
Coefficients to construct the PRS_{313} are shown in Table S3. The PRS_{313} was standardized by SD=0.61, in line with Mavaddat et al[7]. The CBC incidences for each age group were calculated based on incidence data from the Netherlands Cancer Registry[1] and relative risks estimated as described in the Material and Methods. Death was taken into account as competing risk.

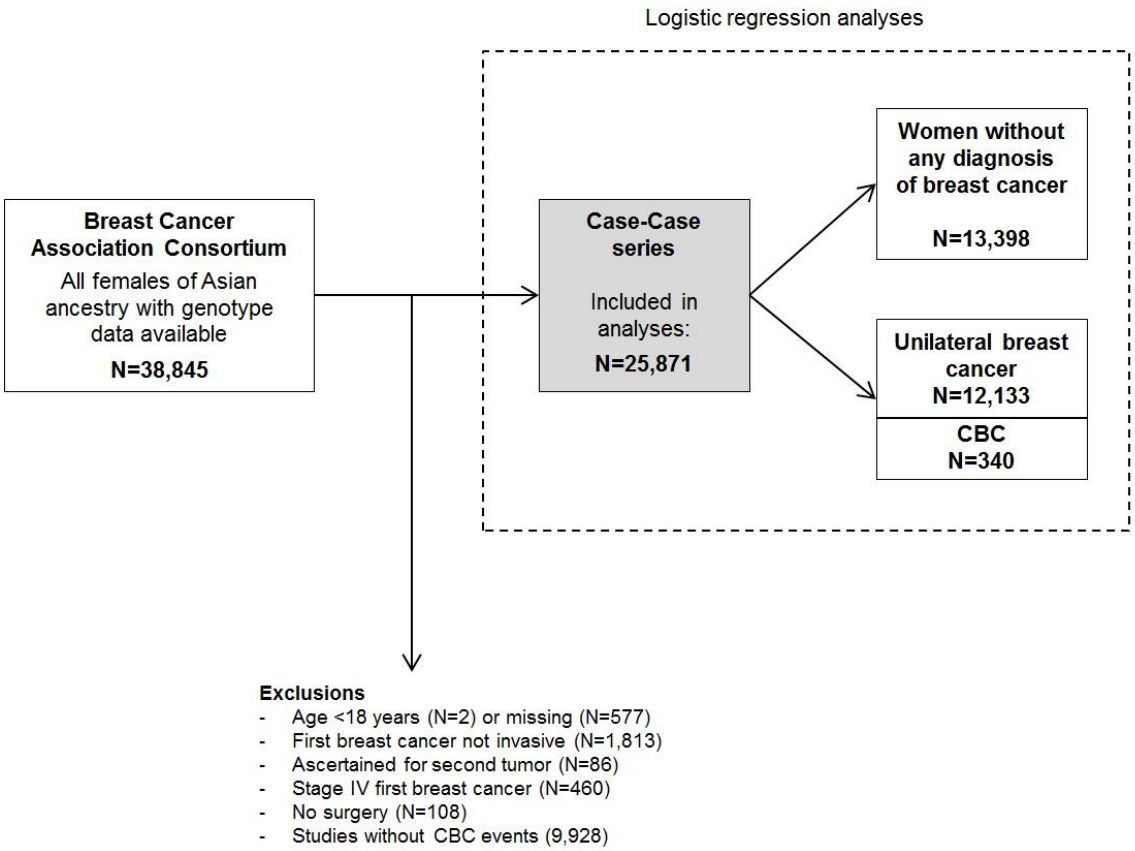
Supplemental Figures

Figure S1A. Overview of the selection of women with breast cancer and control women for the European series



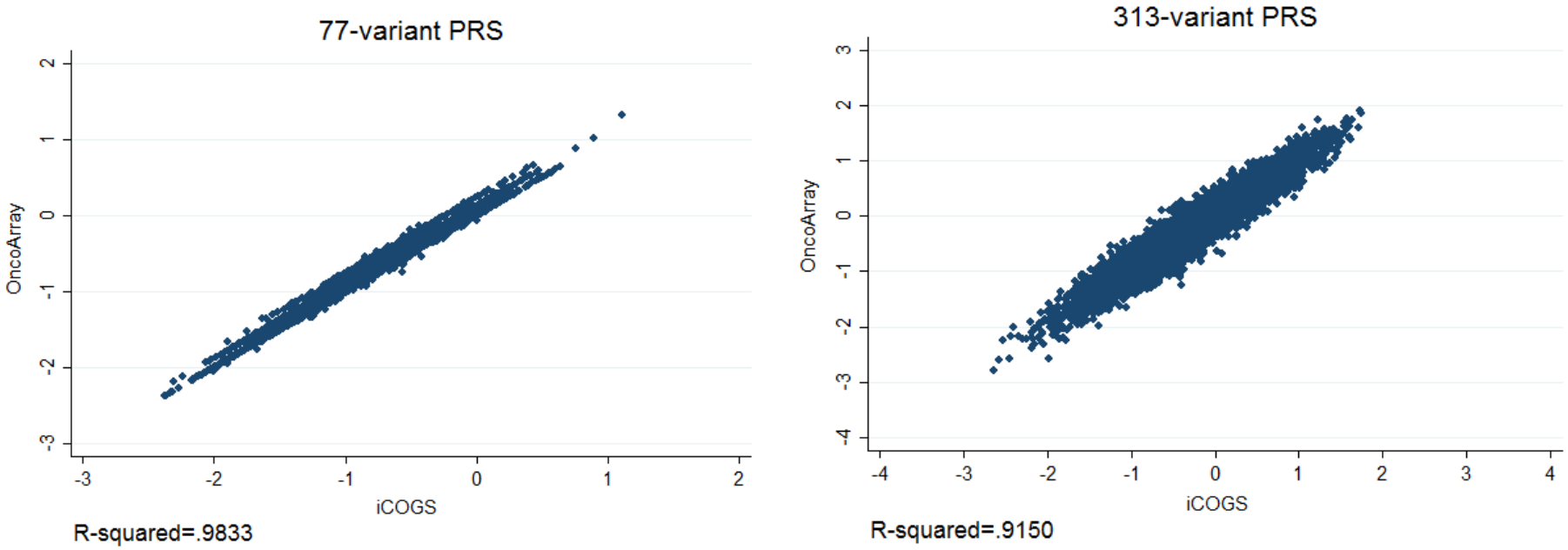
Abbreviations: CBC = contralateral breast cancer
For a complete overview of all studies see Table S1
[a] Excluded studies: CBCS, GLACIER, HMBCS, TNBCC
[b] Excluded studies: BCFR-NY, BCFR-UTAH, CNIO-BCS, DIETCOMPLYF, FHRISK, GESBC, HABCS, HUBCS, ICICLE, KBCP, MCCS, MMHS, NCBCS, PREFACE, SUCCESSB, SUCCESSC
[c] These studies dropped out because for these analyses the definition of CBC is based on the criteria that the CBC was diagnosed at least three months after the first breast cancer diagnosis

Figure S1B. Overview of the selection of women with breast cancer and control women for Asian series



Abbreviations: CBC = contralateral breast cancer

Figure S2. Correlation of total variant scores between the iCOGS array and OncoArray for the 77-variant PRS and the 313-variant PRS[a,b]

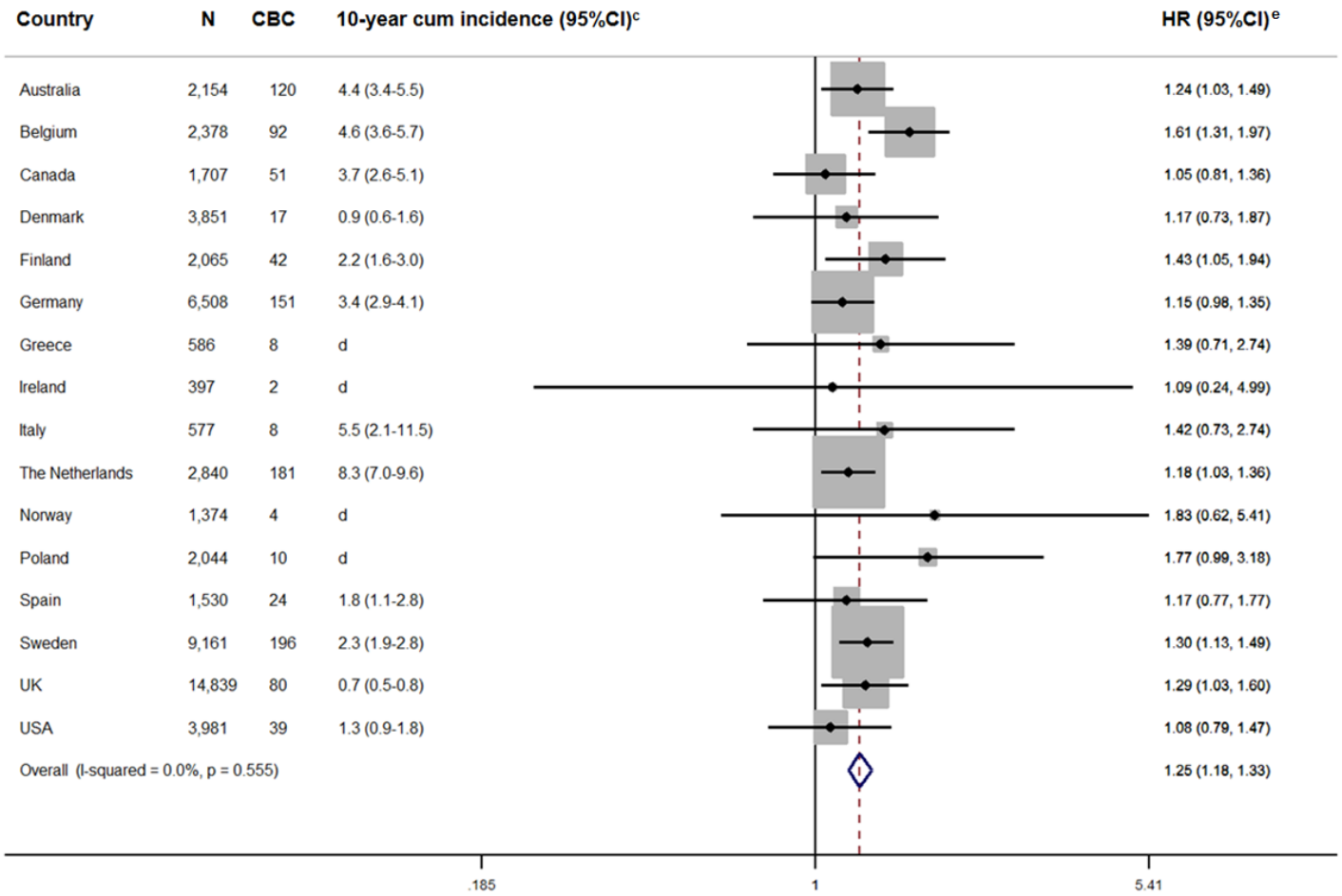


Abbreviations: PRS = polygenic risk score, SD = standard deviation

[a] We evaluated consistency between iCOGS and OncoArray using the intraclass correlation coefficient (ICC), showing a ICC of 0.99 (95%CI=0.99-0.99) for the PRS_{77}, and an ICC of 0.96 (95%CI=0.95-0.96) for the PRS_{313}, based on N=9,071 observations

[b] Coefficients to construct the PRSs are shown in Table S3. The PRSs were standardized by the same SD as was used by Mavaddat et al.[1]. The SD was 0.45 for overall breast cancer PRS_{77}, and 0.61 for overall breast cancer PRS_{313}

Figure S3. Forest plot of the association between the 313-variant PRS and contralateral breast cancer risk by country[a,b]



Abbreviations: PRS = polygenic risk score, N = number of women, CBC = contralateral breast cancer, cum = cumulative, CI = confidence interval, HR = hazard ratio, SD = standard deviation

Fixed effect meta-analysis was used to calculate I-squared and P-value for heterogeneity

[a] Republic of North Macedonia was left out this plot because of a too small sample size (N=76 women including N=2 CBC events)

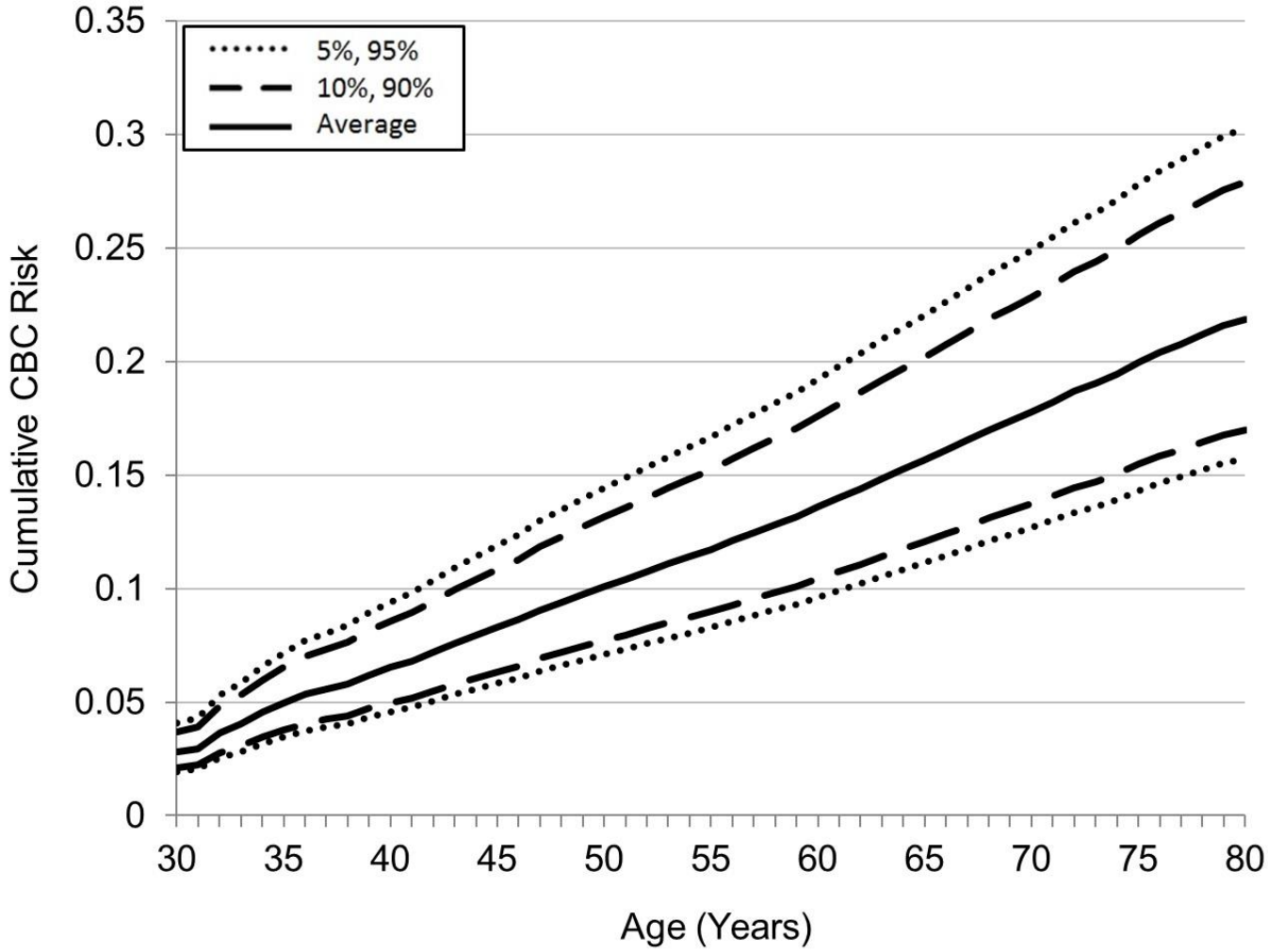
[b] Coefficients to construct the PRS_{313} are shown in Table S3. The PRS_{313} was standardized by SD=0.61, in line with Mavaddat et al.[1]

[c] The 10-year cumulative incidence of CBC was estimated with time since first breast cancer as time scale, and distant metastases (where available) and death as competing risks

[d] Follow-up too short for calculating 10-year cumulative incidence

[e] HR per SD. The analyses were performed with attained age as the time scale

Figure S4. Predicted contralateral breast cancer risk by percentile of the 313-variant PRS (PRS_{313})



Abbreviations: PRS = polygenic risk score, CBC = contralateral breast cancer

Coefficients to construct the PRS_{313} are shown in Table S3. The PRS_{313} was standardized by SD=0.61, in line with Mavaddat et al[1]. The CBC incidences were calculated based on incidence data from the Netherlands Cancer Registry[2] and relative risks estimated as described in the Material and Methods. In contrast to Figure 2, death was not taken into account as competing risk.

Supplemental Tables

Table S1. Study characteristics of included studies of the Breast Cancer Association Consortium

Table S2. Studies and samples included in the analyses using the case-case series, cohort, and validation set

	European								Asian		
	Case-case series N studies = 62			Cohort N studies = 42		Validation set N studies = 24			Case-case series N studies = 8		
Studies	Control women[a]	Unilateral BC	CBC	Unilateral BC	CBC	Control women[a]	Unilateral BC	CBC	Control women[a]	Unilateral BC	CBC
ABCFS	738	1,149	127	1,021	93	-	-	-	-	-	-
ABCS	1,567	1,047	54	519	14	-	-	-	-	-	-
ABCS-F	0	861	91	363	17	-	-	-	-	-	-
ABCTB	375	900	17	708	1	74	180	8	-	-	-
BBCC	711	845	58	766	6	49	56	5	-	-	-
BBCS	1,768	1,266	80	466	1	-	-	-	-	-	-
BCEES	-	-	-	-	-	166	133	0	-	-	-
BCFR-NY	27	340	61	-	-	-	-	-	-	-	-
BCFR-PA	0	104	14	69	4	-	-	-	-	-	-
BCFR-UTAH	0	13	87	-	-	-	-	-	-	-	-
BCINIS	-	-	-	-	-	144	262	0	-	-	-
BIGGS	49	713	50	395	2	-	-	-	-	-	-
BREOGAN	725	1,245	19	1,233	15	145	238	4	-	-	-
BSUCH	1,122	900	36	727	3	-	-	-	-	-	-
CBCS	817	530	21	-	-	163	105	4	170	238	10
CCGP	321	598	19	578	8	66	125	7	-	-	-
CGPS	5,250	4,135	60	3,834	17	142	227	3	-	-	-
CNIO-BCS	829	742	5	-	-	-	-	-	-	-	-
CTS	-	-	-	-	-	115	220	0	-	-	-
DIETCOMPLYF	0	704	1	-	-	-	-	-	-	-	-
FHRISK	0	119	2	-	-	-	-	-	-	-	-
GC-HBOC	1,732	2,690	230	1,406	47	-	-	-	-	-	-
GENICA	711	869	26	869	1	56	89	2	-	-	-
GESBC	181	303	3	-	-	-	-	-	-	-	-
GLACIER	0	1,733	230	-	-	-	-	-	-	-	-
HABCS	863	774	84	-	-	173	141	6	-	-	-
HCSC	0	362	13	273	9	-	-	-	-	-	-
HEBCS	1,060	1,632	116	1,578	41	-	-	-	-	-	-
HERPACC	-	-	-	-	-	-	-	-	1,659	756	18
HKBCS	-	-	-	-	-	-	-	-	451	403	12
HMBCS	345	729	28	-	-	-	-	-	-	-	-
HUBCS	116	198	2	-	-	-	-	-	-	-	-
ICICLE	1	138	12	-	-	-	-	-	-	-	-
KARBAC	0	761	46	443	32	-	-	-	-	-	-
KARMA	5,981	2,314	96	2,188	33	597	185	10	-	-	-
KBCP	431	516	9	-	-	-	-	-	-	-	-
KCONFAB/AOCS	898	397	83	305	26	-	-	-	-	-	-
LMBC	1,821	3,016	208	2,286	92	87	142	14	-	-	-
MABCS	88	80	9	74	2	-	-	-	-	-	-

MARIE	2,066	1,540	115	1,535	53	-	-	-	-	-	-
MBCSG	766	1,015	150	569	8				-	-	-
MCBCS	2,093	1,999	59	1,903	6	35	96	3	-	-	-
MCCS	1,207	1,034	2	-	-	142	86	0	-	-	-
MEC	1,123	1,016	38	988	23	-	-	-	-	-	-
MISS	1,529	582	6	563	3	304	83	0	-	-	-
MMHS	1,635	273	4	-	-	320	48	4	-	-	-
MYBRCA	-	-	-	-	-	-	-	-	4,197	3,652	105
NBCS	212	2,334	31	1,370	4	-	-	-	-	-	-
NBHS	-	-	-	-	-	122	79	0	-	-	-
NC-BCFR	150	614	69	602	5	-	-	-	52	391	33
NCBCS	1,006	1,988	42	-	-	-	-	-	-	-	-
OBCS	414	467	10	445	1	-	-	-	-	-	-
OFBCR	728	1,908	143	1,656	51	-	-	-	-	-	-
ORIGO	0	1,090	89	1,053	69	132	134	15	-	-	-
PBCS	2,082	1,719	40	1,625	9	331	215	2	-	-	-
PKARMA	5,435	4,81	277	4,685	124	1	4	0	-	-	-
POSH	0	1,069	19	1,063	16	-	-	-	-	-	-
PREFACE	0	2,73	90	-	-	-	-	-	-	-	-
PROCAS	1,647	488	9	422	3	-	-	-	-	-	-
RBCS	0	873	152	724	81	-	-	-	-	-	-
SASBAC	1,378	1,118	22	1,086	5	-	-	-	-	-	-
SBCS	848	748	14	691	1	-	-	-	-	-	-
SEARCH	9,056	12,423	118	12,117	59	197	628	0	-	-	-
SEBCS	-	-	-	-	-	-	-	-	2,236	2,080	21
SGBCC	-	-	-	-	-	-	-	-	4,141	1,250	124
SKKDKFZS	29	1,084	71	1,054	41	-	-	-	-	-	-
SMC	-	-	-	-	-	141	244	0	-	-	-
SUCCESSB	0	438	2	-	-	-	-	-	-	-	-
SUCCESSC	0	2,807	29	-	-	-	-	-	-	-	-
SZBCS	489	676	6	409	1	-	-	-	-	-	-
TNBCC	152	1,037	2	-	-	-	-	-	-	-	-
TWBCS	-	-	-	-	-	-	-	-	492	1,250	17
UCIBCS	258	397	1	380	1	51	61	7	-	-	-
Total	**62,830**	**81,000**	**3,607**	**55,041**	**1,027**	**3,753**	**3,781**	**94**	**13,398**	**12,133**	**340**
Characteristics											
Invasiveness in situ	-	excluded	361	excluded	104	-	3[b]	7	-	excluded	67
invasive	-	79,876	2,200	54,675	670	-	3,777	60	-	11,929	209
unknown	-	1,124	1,046	366	253	-	1	27	-	204	64
ER status negative	-	13,828	446	9,333	105	-	766	8	-	3,457	54
positive	-	52,238	2,048	37,420	289	-	3,001	47	-	7,826	163
unknown	-	14,934	1,113	8,288	633	-	14	39	-	850	123

Abbreviations: BC = breast cancer, CBC = contralateral breast cancer, ER = estrogen receptor

[a] Without any diagnosis of breast cancer

[b] Due to the use of a new freeze of the BCAC data, N=3 breast cancers were now defined as in situ, which had previously been defined as invasive; the original validation dataset contained data of two additional studies[1]

Table S3. Variant information and breast cancer risk coefficients for the 77-variant PRS, 313-variant PRS, and ER-specific PRSs; previously published in Mavaddat et al.[1; 3]

Table S4. Patient, tumor, and treatment characteristics of all women diagnosed with first invasive breast cancer since 1990 (European cohort)

Characteristics	Number of women (%)[a]
Total	56,068 (100)
Median age at first diagnosis in years (range)	56 (18-98)
Year of diagnosis	
1990-1994	3,029 (5.4)
1995-1999	10,153 (18.1)
2000-2004	18,484 (33.0)
2005-2009	17,575 (31.3)
2010-2015	6,827 (12.2)
Family history (first degree relative)	
no	33,623 (76.4)
yes	10,369 (23.6)
unknown	12,076
Nodal status	
negative	29,070 (61.9)
positive	17,903 (38.1)
unknown	9,095
Tumor size, cm	
≤2	28,057 (63.8)
(2, 5]	14,138 (32.2)
>5	1,750 (4.0)
unknown	12,123
Differentiation grade	
I	8,721 (19.5)
II	21,621 (48.3)
III	14,454 (32.3)
unknown	11,272
Morphology	
ductal	37,324 (76.6)
lobular	5,878 (12.1)
mixed (ductal and lobular)	2,174 (4.5)
other	3,344 (6.9)
unknown	7,348

ER-status	
negative	9,527 (20.0)
positive	38,090 (80.0)
unknown	8,451
PR-status	
negative	13,098 (32.6)
positive	27,044 (67.4)
unknown	15,926
HER2-status	
negative	23,787 (82.7)
positive	4,969 (17.3)
unknown	27,312
Surgery	
yes, breast saving	16,468 (42.3)
yes, mastectomy	11,315 (29.1)
yes, type unknown	11,163 (28.7)
unknown	17,122
(Neo)adjuvant chemotherapy	
no	18,110 (49.4)
yes	18,559 (50.6)
unknown	19,399
(Neo)adjuvant endocrine therapy	
no	10,781 (28.3)
yes	27,322 (71.7)
unknown	17,965
Radiotherapy	
no	11,023 (27.4)
yes	29,142 (72.6)
unknown	15,903

Abbreviations: ER = estrogen receptor, PR = progesterone receptor, HER2 = human epidermal growth factor receptor 2
[a] Total may not be 100% because of rounding

Table S5. Association between the 313-variant PRS (PRS_{313}) and contralateral breast cancer risk in the European cohort

Percentile categories of the PRS_{313}	No. of women	No. of CBC	HR per unit SD[a]	95%CI	P-value
0^{th} to 10^{th}	5,607	65	0.59	0.45-0.78	<.001
10^{th} to 20^{th}	5,606	79	0.71	0.55-0.92	.01
20^{th} to 40^{th}	11,214	165	0.74	0.60-0.90	.003
40^{th} to 60^{th}	11,214	224	1.00	Ref.	-
60^{th} to 80^{th}	11,214	208	0.90	0.74-1.08	.25
80^{th} to 90^{th}	5,607	121	1.05	0.84-1.31	.69
90^{th} to 100^{th}	5,606	165	1.38	1.13-1.69	.002

Abbreviations: PRS = polygenic risk score, No = number, CBC = contralateral breast cancer, HR = hazard ratio, CI = confidence interval, SD = standard deviation
[a] The analysis was performed with attained age as time scale. Coefficients to construct the PRS_{313} are shown in Table S3. The PRS_{313} was standardized by SD=0.61, in line with Mavaddat et al.[1]

Table S6. Multivariable Cox regression models of contralateral breast cancer risk by 313-variant PRS (PRS_{313}) in all women, all women excluding studies oversampling cases with family history, and those with complete covariate information

		All patients N=56,068 (CBC=1,027)			All women excluding studies oversampling cases with family history N=51,883 (CBC=829)			Complete case N=12,065 (CBC=193)		
		HR per unit SD[a]	**95%CI**	**P-value**	**HR per unit SD[a]**	**95%CI**	**P-value**	**HR per unit SD[a]**	**95%CI**	**P-value**
Model 1										
PRS_{313}[b]		1.25	1.18-1.33	<.001	1.26	1.17-1.34	<.001	1.35	1.17-1.56	<.001
Model 2										
PRS_{313}[b]		1.23	1.16-1.31	<.001	1.25	1.17-1.34	<.001	1.33	1.15-1.54	<.001
Family history	yes vs. no	1.43	1.24-1.64	<.001	1.34	1.13-1.59	.001	1.49	1.06-2.09	.02
	unknown vs. no	0.93	0.75-0.16	.54	0.92	0.73-1.16	.47	-	-	-
Model 3										
PRS_{313}[b]		1.25	1.18-1.33	<.001	1.26	1.17-1.34	<.001	1.35	1.17-1.56	<.001
Nodal status	positive vs. negative	1.05	0.91-1.20	.50	1.07	0.92-1.25	.37	1.14	0.85-1.53	.37
	unknown vs. no	1.26	1.04-1.53	.02	1.29	1.04-1.60	.02	-	-	-
Model 4										
PRS_{313}[b]		1.25	1.18-1.33	<.001	1.26	1.18-1.35	<.001	1.35	1.17-1.56	<.001
Tumor size,	(2-5] vs. ≤2	1.08	0.92-1.25	.34	1.12	0.95-1.32	.20	0.93	0.68-1.27	.66
	>5 vs. ≤2	1.37	0.99-1.89	.06	1.45	1.02-2.07	.04	1.63	0.93-2.85	.09
	unknown vs. ≤2	1.23	1.04-1.47	.02	1.14	0.94-1.39	.18	-	-	-
Model 5										
PRS_{313}[b]		1.25	1.17-1.33	<.001	1.25	1.17-1.34	<.001	1.35	1.17-1.57	<.001
Differentiation grade	II vs. I	0.93	0.76-1.13	.45	0.99	0.80-1.24	.94	0.98	0.65-1.48	.93
	III vs. I	0.90	0.73-1.12	.35	0.97	0.76-1.24	.81	1.09	0.70-1.69	.69
	unknown vs. I	1.20	0.96-1.49	.11	1.45	1.13-1.86	.004	-	-	-
Model 6										
PRS_{313}[b]		1.25	1.17-1.33	<.001	1.25	1.17-1.34	<.001	1.33	1.16-1.54	<.001
Morphology	lobular vs. ductal	1.26	1.03-1.53	.03	1.34	1.08-1.67	.008	1.48	0.99-2.21	.05
	mixed (ductal and lobular) vs. ductal	1.28	0.94-1.73	.11	1.36	0.98-1.88	.06	1.48	0.87-2.54	.15
	other vs. ductal	1.04	0.81-1.33	.75	0.91	0.66-1.24	.55	1.24	0.69-2.21	.47
	unknown vs. ductal	1.77	1.42-2.19	<.001	1.82	1.44-2.30	<.001	-	-	-
Model 7										
PRS_{313}[b]		1.25	1.18-1.33	<.001	1.26	1.18-1.35	<.001	1.35	1.17-1.56	<.001
ER-status	positive vs. negative	0.88	0.75-1.04	.14	0.86	0.72-1.03	.11	0.90	0.62-1.32	.60
	unknown vs. negative	1.16	0.93-0.43	.19	1.11	0.86-1.43	.43	-	-	-
Model 7										
PRS_{313}[b]		1.25	1.18-1.33	<.001	1.26	1.18-1.35	<.001	1.35	1.17-1.56	<.001

PR-status	positive vs. negative	0.95	0.81-1.11	.51	0.92	0.78-1.09	.32	0.91	0.66-1.25	.56
	unknown vs. negative	1.15	0.95-1.40	.14	1.10	0.88-1.37	.40	-	-	-
Model 9										
PRS_{313}[b]		1.25	1.18-1.33	<.001	1.26	1.17-1.34	<.001	1.34	1.16-1.55	<.001
HER2-status	positive vs. negative	0.84	0.64-1.11	.22	0.76	0.56-1.05	.10	0.70	0.45-1.10	.12
	unknown vs. negative	1.29	1.11-1.50	.001	1.28	1.08-1.52	.004	-	-	-
Model 10										
PRS_{313}[b]		1.25	1.17-1.32	<.001	1.26	1.17-1.34	<.001	1.35	1.16-1.56	<.001
Chemotherapy	yes vs. no	0.86	0.73-1.01	.06	0.99	0.83-1.19	.92	0.89	0.64-1.25	.51
	unknown vs. no	1.09	0.91-1.31	.34	1.20	0.97-1.47	.09	-	-	-
Model 11										
PRS_{313}[b]		1.25	1.18-1.33	<.001	1.26	1.18-1.35	<.001	1.36	1.17-1.57	<.001
Endocrine therapy	yes vs. no	0.75	0.64-0.88	.001	0.92	0.75-1.12	.41	0.78	0.55-1.11	.17
	unknown vs. no	0.90	0.75-1.09	.28	1.11	0.87-1.41	.39	-	-	-
Model 12										
PRS_{313}[b]		1.25	1.17-1.32	<.001	1.26	1.17-1.34	<.001	1.35	1.17-1.56	<.001
Radiotherapy	yes vs. no	1.00	0.85-1.18	1.00	0.98	0.82-1.18	.85	1.35	0.88-2.08	.17
	unknown vs. no	1.41	1.14-1.74	.001	1.18	0.93-1.50	.17	-	-	-
Model 13										
PRS_{313}[b]		1.25	1.17-1.32	<.001	1.25	1.17-1.34	<.001	1.34	1.16-1.55	<.001
Year of first breast cancer diagnosis		0.95	0.94-0.96	<.001	0.95	0.93-0.96	<.001	0.90	0.86-0.95	<.001
Model 14										
PRS_{313}[b]	full model[c]	1.23	1.16-1.31	<.001	1.25	1.16-1.33	<.001	1.33	1.15-1.53	<.001

Abbreviations: PRS = polygenic risk score, CBC = contralateral breast cancer, HR = hazard ratio, CI = confidence interval, SD = standard deviation, ER = estrogen receptor, PR = progesterone receptor, HER2 = human epidermal growth factor receptor 2

[a] All analyses were performed with attained age as the time scale

[b] Coefficients to construct the PRS_{313} are shown in Table S3. The PRS_{313} was standardized by SD=0.61, in line with Mavaddat et al.[1]

[c] Adjusted for family history, nodal status, tumor size, differentiation grade, morphology, ER status, HER2 status, chemotherapy, endocrine therapy, radiotherapy, and year of first breast cancer diagnosis

Table S7. Five- and ten-year cumulative risks of contralateral breast cancer by the 313-variant PRS (PRS_{313}) for different age groups

	5-year cumulative CBC risks (%) range by age					10-year cumulative CBC risks (%) range by age				
Age at first breast cancer diagnosis (years)	**5^{th} percentile PRS_{313}**	**10^{th} percentile PRS_{313}**	**50^{th} percentile PRS_{313}**	**90^{th} percentile PRS_{313}**	**95^{th} percentile PRS_{313}**	**5^{th} percentile PRS_{313}**	**10^{th} percentile PRS_{313}**	**50^{th} percentile PRS_{313}**	**90^{th} percentile PRS_{313}**	**95^{th} percentile PRS_{313}**
30-34	1.9-3.3	2.1-3.6	2.8-4.7	3.7-6.2	4.0-6.8	3.3-4.4	3.6-4.8	4.7-6.3	6.2-8.3	6.8-9.1
35-39	0.8-2.2	0.9-2.4	1.2-3.2	1.6-4.2	1.7-4.6	2.2-3.9	2.4-4.2	3.2-5.5	4.2-7.2	4.6-8.0
40-44	1.5-2.9	1.7-3.2	2.2-4.2	2.9-5.5	3.2-6.0	2.9-4.9	3.2-5.3	4.2-7.0	5.5-9.1	6.0-10.0
45-49	1.4-2.5	1.5-2.8	2.0-3.7	2.6-4.8	2.9-5.3	2.5-4.2	2.8-4.5	3.7-6.0	4.8-7.8	5.3-8.6
50-54	1.4-2.9	1.5-3.1	2.0-4.1	2.6-5.5	2.9-6.0	2.9-4.8	3.1-5.3	4.1-6.9	5.5-9.1	6.0-10.0
55-59	1.6-3.3	1.8-3.6	2.4-4.7	3.1-6.2	3.4-6.8	3.3-5.3	3.6-5.7	4.7-7.5	6.2-9.8	6.8-10.8
60-64	1.8-3.5	1.9-3.8	2.6-5.0	3.4-6.5	3.7-7.2	3.5-5.5	3.8-6.0	5.0-7.9	6.5-10.3	7.2-11.3
65-70	1.5-3.5	1.7-3.8	2.2-5.0	2.9-6.6	3.2-7.2	3.5-4.6	3.8-5.0	5.0-6.6	6.6-8.7	7.2-9.5

Abbreviations: PRS = polygenic risk score, CBC = contralateral breast cancer

Coefficients to construct the PRS_{313} are shown in Table S3. The PRS_{313} was standardized by SD=0.61, in line with Mavaddat et al[1]. The CBC incidences for each age group were calculated based on incidence data from the Netherlands Cancer Registry[2] and relative risks estimated as described in the Material and Methods. In contrast to Table 4, death was not taken into account as competing risk

Table S8. Estimates of unilateral- and contralateral breast cancer risk by the 313-variant PRS (PRS_{313}) in the European case-case series and the Asian case-case series

	European						Asian		
	Case-case series[a]			Validation set[b]			Case-case series[a]		
PRS_{313}[c]	**OR per unit SD**	**95%CI**	**P-value**	**OR per unit SD**	**95%CI**	**P-value**	**OR per unit SD**	**95%CI**	**P-value**
Unilateral breast cancer versus control	1.82	1.80-1.84	<.001	1.67	1.59-1.76	<.001	1.56	1.52-1.60	<.001
CBC versus unilateral breast cancer	1.30	1.26-1.35	<.001	1.39	1.13-1.70	.002	1.15	1.02-1.29	.02

Abbreviations: PRS = polygenic risk score, CBC = contralateral breast cancer, OR = odds ratio, SD = standard deviation, CI = confidence interval

[a] Adjusted for country and age. For all women with unilateral- and contralateral breast cancer we used age at first breast cancer diagnosis, and for control women without any diagnosis of breast cancer we used age at baseline questionnaire.

[b] The validation set was previously used to develop the PRS_{313}; see details in materials and methods. For analyses in the current paper, this set is nested within the case-case series. These analyses were additionally adjusted for 10 principal components for comparability with the originally published PRS_{313} overall estimates[1]

[c] Coefficients to construct the PRS_{313} are shown in Table S3. The PRS_{313} was standardized by SD=0.61, in line with Mavaddat et al.[1]

Supplemental Note

Our initial aim was to externally validate our results using the UK Biobank, which seemed the most suitable cohort given the large number of women diagnosed with breast cancer with information available on the PRS_{313},. However, when we started the analyses, it turned out that the UK Biobank had no information available on the laterality of the second breast tumor. Therefore, we were unable to distinguish between ipsilateral and contralateral breast cancer, and had to define our endpoint in these analyses as 'any second breast cancer'. In addition, in comparison to our analyses in the BCAC, we were unable to exclude patients diagnosed with stage IV invasive first breast cancer from the UK Biobank cohort, and had limited information on metastases developed during follow-up.

The association between the overall breast cancer PRS_{313} and (any) second breast cancer was evaluated among women aged ≥18 years of European ancestry from the UK Biobank cohort who had had a diagnosis of invasive first breast cancer. UK Biobank samples were genotyped using Affymetrix UK BiLEVE Axiom array and Affymetrix UK Biobank Axiom® array and imputed to the combined 1000 Genome Project v3 and UK10K reference panels using SHAPEIT3 and IMPUTE3[4]. The lowest imputation info score for the variants used in these analyses was 0.86. Samples were included for this analysis of the UK BIOBANK study on the basis of female sex (genetic and self-reported) and ethnicity filter (Europeans/White British ancestry subset). Duplicates and individuals with high degree of relatedness (samples which have >10 putative third degree relatives) were removed, and we randomly excluded one of each related pair first-degree relatives. Samples were also excluded on standard quality control criteria. The PRS_{313} was calculated as a weighted sum of the minor allele dosages; the variant selection and weights are as given by Mavaddat et al[1]. The PRS_{313} was standardized by SD=0.61, in line with our BCAC analyses and Mavaddat et al[1].

The final cohort included 10,567 women with invasive breast cancer among whom 302 registry-confirmed second breast cancers developed over 59,260 person-years of follow-up. A Cox proportional hazards model was used to assess the association between PRS_{313} and second breast cancer risk. Time at risk started three months after the age of first breast cancer diagnosis, where this was diagnosed after the baseline questionnaire date, or three months after the baseline questionnaire where first breast cancer was diagnosed before the baseline questionnaire date. Time at risk ended at the age of second breast cancer diagnosis (ipsilateral or contralateral), distant metastasis (where available), death or end of follow-up (at latest December 10, 2016). Potential effect modification of the PRS_{313} by age was evaluated by adding an interaction term (PRS_{313} x age at first breast cancer diagnosis [continuous]) in the model. We performed a separate analysis for invasive second breast cancer (241 breast cancers), where we censored on in situ second breast cancer.

The HR for a second breast cancer (in situ or invasive) per SD of PRS_{313} in the UK Biobank cohort was 1.13 (95%CI=1.01-1.26). We found no indication for interaction with age at first breast cancer diagnosis ($HR_{interaction}$=1.00, 95%CI=0.99-1.01; P=0.87). When analyses were restricted to invasive second breast cancer, the HR per SD was 1.13 (95%CI=1.00-1.29).

Supplemental Acknowledgements

Funding

This work was supported by the Alpe d'HuZes/Dutch Cancer Society (KWF Kankerbestrijding) (grant number A6C/6253).

BCAC is funded by Cancer Research UK [C1287/A16563, C1287/A10118], the European Union's Horizon 2020 Research and Innovation Programme (grant numbers 634935 and 633784 for BRIDGES and B-CAST respectively), and by the European Community´s Seventh Framework Programme under grant agreement number 223175 (grant number HEALTH-F2-2009-223175) (COGS). The EU Horizon 2020 Research and Innovation Programme funding source had no role in study design, data collection, data analysis, data interpretation or writing of the report.

Genotyping of the OncoArray was funded by the NIH Grant U19 CA148065, and Cancer UK Grant C1287/A16563 and the PERSPECTIVE project supported by the Government of Canada through Genome Canada and the Canadian Institutes of Health Research (grant GPH-129344) and, the Ministère de l'Économie, Science et Innovation du Québec through Genome Québec and the PSRSIIRI-701 grant, and the Quebec Breast Cancer Foundation. Funding for the iCOGS infrastructure came from: the European Community's Seventh Framework Programme under grant agreement n° 223175 (HEALTH-F2-2009-223175) (COGS), Cancer Research UK (C1287/A10118, C1287/A10710, C12292/A11174, C1281/A12014, C5047/A8384, C5047/A15007, C5047/A10692, C8197/A16565), the National Institutes of Health (CA128978) and Post-Cancer GWAS initiative (1U19 CA148537, 1U19 CA148065 and 1U19 CA148112 - the GAME-ON initiative), the Department of Defence (W81XWH-10-1-0341), the Canadian Institutes of Health Research (CIHR) for the CIHR Team in Familial Risks of Breast Cancer, and Komen Foundation for the Cure, the Breast Cancer Research Foundation, and the Ovarian Cancer Research Fund.

The Australian Breast Cancer Family Study (ABCFS) was supported by grant UM1 CA164920 from the National Cancer Institute (USA). The content of this manuscript does not necessarily reflect the views or policies of the National Cancer Institute or any of the collaborating centers in the Breast Cancer Family Registry (BCFR), nor does mention of trade names, commercial products, or organizations imply endorsement by the USA Government or the BCFR. The ABCFS was also supported by the National Health and Medical Research Council of Australia, the New South Wales Cancer Council, the Victorian Health Promotion Foundation (Australia) and the Victorian Breast Cancer Research Consortium. J.L.H. is a National Health and Medical Research Council (NHMRC) Senior Principal Research Fellow. M.C.S. is a NHMRC Senior Research Fellow. The ABCS study was supported by the Dutch Cancer Society [grants NKI 2007-3839; 2009 4363]. The Australian Breast Cancer Tissue Bank (ABCTB) was supported by the National Health and Medical Research Council of Australia, The Cancer Institute NSW and the National Breast Cancer Foundation. The work of the BBCC was partly funded by ELAN-Fond of the University Hospital of Erlangen. The BBCS is funded by Cancer Research UK and Breast Cancer Now and acknowledges NHS funding to the NIHR Biomedical Research Centre, and the National Cancer Research Network (NCRN). The BCEES was funded by the National Health and Medical Research Council, Australia and the Cancer Council Western Australia and acknowledges funding from the National Breast Cancer Foundation (JS). For the BCFR-NY, BCFR-PA, BCFR-UT this work was supported by grant UM1 CA164920 from the National Cancer Institute. The content of this manuscript does not necessarily reflect the views or policies of the National Cancer Institute or any of the collaborating centers in the Breast Cancer Family Registry (BCFR), nor does mention of trade names, commercial products, or organizations imply endorsement by the US Government or the BCFR. The BCINIS study was supported in part by the BCRF (Breast Cancer Research Foundation, NY, USA). For BIGGS, ES is supported by NIHR Comprehensive Biomedical Research Centre, Guy's & St. Thomas' NHS Foundation Trust in partnership with King's College London, United Kingdom. IT is supported by the Oxford Biomedical Research Centre. The BREast Oncology GAlician Network

(BREOGAN) is funded by Acción Estratégica de Salud del Instituto de Salud Carlos III FIS PI12/02125/Cofinanciado FEDER, FIS PI17/00918/Cofinanciado FEDER; Acción Estratégica de Salud del Instituto de Salud Carlos III FIS Intrasalud (PI13/01136); Programa Grupos Emergentes, Cancer Genetics Unit, Instituto de Investigacion Biomedica Galicia Sur. Xerencia de Xestion Integrada de Vigo-SERGAS, Instituto de Salud Carlos III, Spain; Grant 10CSA012E, Consellería de Industria Programa Sectorial de Investigación Aplicada, PEME I + D e I + D Suma del Plan Gallego de Investigación, Desarrollo e Innovación Tecnológica de la Consellería de Industria de la Xunta de Galicia, Spain; Grant EC11-192. Fomento de la Investigación Clínica Independiente, Ministerio de Sanidad, Servicios Sociales e Igualdad, Spain; and Grant FEDER-Innterconecta. Ministerio de Economia y Competitividad, Xunta de Galicia, Spain. The BSUCH study was supported by the Dietmar-Hopp Foundation, the Helmholtz Society and the German Cancer Research Center (DKFZ). CBCS is funded by the Canadian Cancer Society (grant # 313404) and the Canadian Institutes of Health Research. CCGP is supported by funding from the University of Crete. The CGPS was supported by the Chief Physician Johan Boserup and Lise Boserup Fund, the Danish Medical Research Council, and Herlev and Gentofte Hospital. The CNIO-BCS was supported by the Instituto de Salud Carlos III, the Fondo de Investigación Sanitario (PI16/00440 with FEDER funds), and CIBERER (Spanish Network on Rare diseases). The CTS was initially supported by the California Breast Cancer Act of 1993 and the California Breast Cancer Research Fund (contract 97-10500) and is currently funded through the National Institutes of Health (R01 CA77398, UM1 CA164917, and U01 CA199277). Collection of cancer incidence data was supported by the California Department of Public Health as part of the statewide cancer reporting program mandated by California Health and Safety Code Section 103885. HAC receives support from the Lon V Smith Foundation (LVS39420). The University of Westminster curates the DietCompLyf database funded by Against Breast Cancer Registered Charity No. 1121258 and the NCRN. FHRISK is funded from NIHR grant PGfAR 0707-10031. The GC-HBOC (German Consortium of Hereditary Breast and Ovarian Cancer) is supported by the German Cancer Aid (grant no 110837, coordinator: Rita K. Schmutzler, Cologne). This work was

also funded by the European Regional Development Fund and Free State of Saxony, Germany (LIFE - Leipzig Research Centre for Civilization Diseases, project numbers 713-241202, 713-241202, 14505/2470, 14575/2470). The GENICA was funded by the Federal Ministry of Education and Research (BMBF) Germany grants 01KW9975/5, 01KW9976/8, 01KW9977/0 and 01KW0114, the Robert Bosch Foundation, Stuttgart, Deutsches Krebsforschungszentrum (DKFZ), Heidelberg, the Institute for Prevention and Occupational Medicine of the German Social Accident Insurance, Institute of the Ruhr University Bochum (IPA), Bochum, as well as the Department of Internal Medicine, Evangelische Kliniken Bonn gGmbH, Johanniter Krankenhaus, Bonn, Germany. The GESBC was supported by the Deutsche Krebshilfe e. V. [70492] and the German Cancer Research Center (DKFZ). GLACIER was supported by Breast Cancer Now, CRUK and Biomedical Research Centre at Guy's and St Thomas' NHS Foundation Trust and King's College London. The HABCS study was supported by the Claudia von Schilling Foundation for Breast Cancer Research, by the Lower Saxonian Cancer Society, and by the Rudolf Bartling Foundation. The HEBCS was financially supported by the Helsinki University Hospital Research Fund, the Finnish Cancer Society, and the Sigrid Juselius Foundation. The HERPACC was supported by MEXT Kakenhi (No. 170150181 and 26253041) from the Ministry of Education, Science, Sports, Culture and Technology of Japan, by a Grant-in-Aid for the Third Term Comprehensive 10-Year Strategy for Cancer Control from Ministry Health, Labour and Welfare of Japan, by Health and Labour Sciences Research Grants for Research on Applying Health Technology from Ministry Health, Labour and Welfare of Japan, by National Cancer Center Research and Development Fund, and "Practical Research for Innovative Cancer Control (15ck0106177h0001)" from Japan Agency for Medical Research and development, AMED, and Cancer Bio Bank Aichi. The HMBCS was supported by a grant from the Friends of Hannover Medical School and by the Rudolf Bartling Foundation. The HUBCS was supported by a grant from the German Federal Ministry of Research and Education (RUS08/017), and by the Russian Foundation for Basic Research and the Federal Agency for Scientific Organizations for support the Bioresource collections and RFBR grants 14-04-97088, 17-29-06014 and 17-44-020498. ICICLE

was supported by Breast Cancer Now, CRUK and Biomedical Research Centre at Guy's and St Thomas' NHS Foundation Trust and King's College London. Financial support for KARBAC was provided through the regional agreement on medical training and clinical research (ALF) between Stockholm County Council and Karolinska Institutet, the Swedish Cancer Society, The Gustav V Jubilee foundation and Bert von Kantzows foundation. The KARMA study was supported by Märit and Hans Rausings Initiative Against Breast Cancer. The KBCP was financially supported by the special Government Funding (EVO) of Kuopio University Hospital grants, Cancer Fund of North Savo, the Finnish Cancer Organizations, and by the strategic funding of the University of Eastern Finland. kConFab is supported by a grant from the National Breast Cancer Foundation, and previously by the National Health and Medical Research Council (NHMRC), the Queensland Cancer Fund, the Cancer Councils of New South Wales, Victoria, Tasmania and South Australia, and the Cancer Foundation of Western Australia. Financial support for the AOCS was provided by the United States Army Medical Research and Materiel Command [DAMD17-01-1-0729], Cancer Council Victoria, Queensland Cancer Fund, Cancer Council New South Wales, Cancer Council South Australia, The Cancer Foundation of Western Australia, Cancer Council Tasmania and the National Health and Medical Research Council of Australia (NHMRC; 400413, 400281, 199600). G.C.T. and P.W. are supported by the NHMRC. RB was a Cancer Institute NSW Clinical Research Fellow. LMBC is supported by the 'Stichting tegen Kanker'. DL is supported by the FWO. The MABCS study is funded by the Research Centre for Genetic Engineering and Biotechnology "Georgi D. Efremov", MASA. The MARIE study was supported by the Deutsche Krebshilfe e.V. [70-2892-BR I, 106332, 108253, 108419, 110826, 110828], the Hamburg Cancer Society, the German Cancer Research Center (DKFZ) and the Federal Ministry of Education and Research (BMBF) Germany [01KH0402]. MBCSG is supported by grants from the Italian Association for Cancer Research (AIRC) and by funds from the Italian citizens who allocated the 5/1000 share of their tax payment in support of the Fondazione IRCCS Istituto Nazionale Tumori, according to Italian laws (INT-Institutional strategic projects "5x1000"). The MCBCS was supported by the NIH grants CA192393, CA116167,

CA176785 an NIH Specialized Program of Research Excellence (SPORE) in Breast Cancer [CA116201], and the Breast Cancer Research Foundation and a generous gift from the David F. and Margaret T. Grohne Family Foundation. The Melbourne Collaborative Cohort Study (MCCS) cohort recruitment was funded by VicHealth and Cancer Council Victoria. The MCCS was further augmented by Australian National Health and Medical Research Council grants 209057, 396414 and 1074383 and by infrastructure provided by Cancer Council Victoria. Cases and their vital status were ascertained through the Victorian Cancer Registry and the Australian Institute of Health and Welfare, including the National Death Index and the Australian Cancer Database. The Multiethnic Cohort Study (MEC) was funded by NIH grant U01 CA164973. The MEC was support by NIH grants CA63464, CA54281, CA098758, CA132839 and CA164973. The MISS study is supported by funding from ERC-2011-294576 Advanced grant, Swedish Cancer Society, Swedish Research Council, Local hospital funds, Berta Kamprad Foundation, Gunnar Nilsson. The MMHS study was supported by NIH grants CA97396, CA128931, CA116201, CA140286 and CA177150. MYBRCA is funded by research grants from the Malaysian Ministry of Higher Education (UM.C/HlR/MOHE/06) and Cancer Research Malaysia. Genotyping for MyBrCa and SGBCC were supported by grants from Newton-Ungku Omar Fund [grant no: MR/P012930/1] and Wellcome Trust [grant no: v203477/Z/16/Z]. The Malaysian Breast Cancer Genetic Study was established using funds from the Malaysian Ministry of Science, and the Malaysian Ministry of Higher Education High Impact Research Grant [grant no: UM.C/HIR/MOHE/06]. The Malaysian Mammographic Density Study was established using funds raised through the Sime Darby LPGA tournament and the High Impact Research Grant. Additional funding was received from Yayasan Sime Darby, PETRONAS, Estee Lauder Group of Companies and other donors of Cancer Research Malaysia. MYMAMMO is supported by research grants from Yayasan Sime Darby LPGA Tournament and Malaysian Ministry of Higher Education (RP046B-15HTM). The NBCS has received funding from the K.G. Jebsen Centre for Breast Cancer Research; the Research Council of Norway grant 193387/V50 (to A-L Børresen-Dale and V.N. Kristensen) and grant 193387/H10 (to A-L Børresen-Dale and V.N. Kristensen), South Eastern Norway

Health Authority (grant 39346 to A-L Børresen-Dale) and the Norwegian Cancer Society (to A-L Børresen-Dale and V.N. Kristensen). The NBHS was supported by NIH grant R01CA100374. Biological sample preparation was conducted the Survey and Biospecimen Shared Resource, which is supported by P30 CA68485. The Northern California Breast Cancer Family Registry (NC-BCFR) and Ontario Familial Breast Cancer Registry (OFBCR) were supported by grant UM1 CA164920 from the National Cancer Institute (USA). The content of this manuscript does not necessarily reflect the views or policies of the National Cancer Institute or any of the collaborating centers in the Breast Cancer Family Registry (BCFR), nor does mention of trade names, commercial products, or organizations imply endorsement by the USA Government or the BCFR. The Carolina Breast Cancer Study was funded by Komen Foundation, the National Cancer Institute (P50 CA058223, U54 CA156733, U01 CA179715), and the North Carolina University Cancer Research Fund. The OBCS was supported by research grants from the Finnish Cancer Foundation, the Academy of Finland (grant number 250083, 122715 and Center of Excellence grant number 251314), the Finnish Cancer Foundation, the Sigrid Juselius Foundation, the University of Oulu, the University of Oulu Support Foundation and the special Governmental EVO funds for Oulu University Hospital-based research activities. The ORIGO study was supported by the Dutch Cancer Society (RUL 1997-1505) and the Biobanking and Biomolecular Resources Research Infrastructure (BBMRI-NL CP16). The PBCS was funded by Intramural Research Funds of the National Cancer Institute, Department of Health and Human Services, USA. The POSH study is funded by Cancer Research UK (grants C1275/A11699, C1275/C22524, C1275/A19187, C1275/A15956 and Breast Cancer Campaign 2010PR62, 2013PR044. PROCAS is funded from NIHR grant PGfAR 0707-10031. DGE, AH and WN are supported by the all Manchester NIHR Biomedical Research Centre (IS-BRC-1215-20007). The RBCS was funded by the Dutch Cancer Society (DDHK 2004-3124, DDHK 2009-4318). The SASBAC study was supported by funding from the Agency for Science, Technology and Research of Singapore (A*STAR), the US National Institute of Health (NIH) and the Susan G. Komen Breast Cancer Foundation. The SBCS was supported by Sheffield Experimental Cancer Medicine Centre and

Breast Cancer Now Tissue Bank. SEARCH is funded by Cancer Research UK [C490/A10124, C490/A16561] and supported by the UK National Institute for Health Research Biomedical Research Centre at the University of Cambridge. The University of Cambridge has received salary support for PDPP from the NHS in the East of England through the Clinical Academic Reserve. SEBCS was supported by the BRL (Basic Research Laboratory) program through the National Research Foundation of Korea funded by the Ministry of Education, Science and Technology (2012-0000347). SGBCC is supported by the National Research Foundation Singapore (NRF-NRFF2017-02), NUS start-up Grant, National University Cancer Institute Singapore (NCIS) Centre Grant [NMRC/CG/NCIS/2010, NMRC/CG/012/2013, CGAug16M005], Saw Swee Hock School of Public Health Research Programme of Research Seed Funding (Breast Cancer Prevention Program), Asian Breast Cancer Research Fund, and the NMRC Clinician Scientist Award (SI Category) [NMRC/CSA-SI/0015/2017]. Controls from Singapore were recruited by the Singapore Consortium of Cohort Studies-Multi-ethnic cohort (SCCS-MEC), which was funded by the Biomedical Research Council, grant number: 05/1/21/19/425. SKKDKFZS is supported by the DKFZ. The SMC is funded by the Swedish Cancer Foundation and the Swedish Research Council (VR 2017-00644) grant for the Swedish Infrastructure for Medical Population-based Life-course Environmental Research (SIMPLER). The SZBCS was supported by Grant PBZ_KBN_122/P05/2004 and the program of the Minister of Science and Higher Education under the name "Regional Initiative of Excellence" in 2019-2022 project number 002 / RID / 2018/19, amount of financing 12 000 000 PLN. The TWBCS is supported by the Taiwan Biobank project of the Institute of Biomedical Sciences, Academia Sinica, Taiwan. The UCIBCS component of this research was supported by the NIH [CA58860, CA92044] and the Lon V Smith Foundation [LVS39420].

Acknowledgements

We thank all the individuals who took part in these studies and all the researchers, clinicians, technicians and administrative staff who have enabled this work to be carried out. The COGS study would not have been possible without the contributions of the following: Kyriaki Michailidou, Qin Wang (BCAC). ABCFS thank Maggie Angelakos, Judi Maskiell, Gillian Dite. ABCS thanks the Blood bank Sanquin, The Netherlands. ABCTB Investigators: Christine Clarke, Deborah Marsh, Rodney Scott, Robert Baxter, Desmond Yip, Jane Carpenter, Alison Davis, Nirmala Pathmanathan, Peter Simpson, J. Dinny Graham, Mythily Sachchithananthan. Samples are made available to researchers on a non-exclusive basis. BBCS thanks Eileen Williams, Elaine Ryder-Mills, Kara Sargus. BCEES thanks Allyson Thomson, Christobel Saunders, Terry Slevin, BreastScreen Western Australia, Elizabeth Wylie, Rachel Lloyd. The BCINIS study would not have been possible without the contributions of Dr. K. Landsman, Dr. N. Gronich, Dr. A. Flugelman, Dr. W. Saliba, Dr. E. Liani, Dr. I. Cohen, Dr. S. Kalet, Dr. V. Friedman, Dr. O. Barnet of the NICCC in Haifa, and all the contributing family medicine, surgery, pathology and oncology teams in all medical institutes in Northern Israel. BIGGS thanks Niall McInerney, Gabrielle Colleran, Andrew Rowan, Angela Jones. The BREOGAN study would not have been possible without the contributions of the following: Manuela Gago-Dominguez, Jose Esteban Castelao, Angel Carracedo, Victor Muñoz Garzón, Alejandro Novo Domínguez, Maria Elena Martinez, Sara Miranda Ponte, Carmen Redondo Marey, Maite Peña Fernández, Manuel Enguix Castelo, Maria Torres, Manuel Calaza (BREOGAN), José Antúnez, Máximo Fraga and the staff of the Department of Pathology and Biobank of the University Hospital Complex of Santiago-CHUS, Instituto de Investigación Sanitaria de Santiago, IDIS, Xerencia de Xestion Integrada de Santiago-SERGAS; Joaquín González-Carreró and the staff of the Department of Pathology and Biobank of University Hospital Complex of Vigo, Instituto de Investigacion Biomedica Galicia Sur, SERGAS, Vigo, Spain. BSUCH thanks Peter Bugert, Medical Faculty Mannheim. CBCS thanks study participants, co-investigators, collaborators and staff of the Canadian Breast Cancer Study, and project coordinators Agnes Lai and

Celine Morissette. CCGP thanks Styliani Apostolaki, Anna Margiolaki, Georgios Nintos, Maria Perraki, Georgia Saloustrou, Georgia Sevastaki, Konstantinos Pompodakis. CGPS thanks staff and participants of the Copenhagen General Population Study. For the excellent technical assistance: Dorthe Uldall Andersen, Maria Birna Arnadottir, Anne Bank, Dorthe Kjeldgård Hansen. The Danish Cancer Biobank is acknowledged for providing infrastructure for the collection of blood samples for the cases. CNIO-BCS thanks Guillermo Pita, Charo Alonso, Nuria Álvarez, Pilar Zamora, Primitiva Menendez, the Human Genotyping-CEGEN Unit (CNIO). The CTS Steering Committee includes Leslie Bernstein, Susan Neuhausen, James Lacey, Sophia Wang, Huiyan Ma, and Jessica Clague DeHart at the Beckman Research Institute of City of Hope, Dennis Deapen, Rich Pinder, and Eunjung Lee at the University of Southern California, Pam Horn-Ross, Peggy Reynolds, Christina Clarke Dur and David Nelson at the Cancer Prevention Institute of California, Hoda Anton-Culver, Argyrios Ziogas, and Hannah Park at the University of California Irvine, and Fred Schumacher at Case Western University. DIETCOMPLYF thanks the patients, nurses and clinical staff involved in the study. The DietCompLyf study was funded by the charity Against Breast Cancer (Registered Charity Number 1121258) and the NCRN. FHRISK thanks NIHR for funding. GC-HBOC thanks Stefanie Engert, Heide Hellebrand, Sandra Kröber and LIFE - Leipzig Research Centre for Civilization Diseases (Markus Loeffler, Joachim Thiery, Matthias Nüchter, Ronny Baber). The GENICA Network: Dr. Margarete Fischer-Bosch-Institute of Clinical Pharmacology, Stuttgart, and University of Tübingen, Germany [HB, Wing-Yee Lo], German Cancer Consortium (DKTK) and German Cancer Research Center (DKFZ), Partner Site Tübingen, 72074 Tübingen, Germany [HB], gefördert durch die Deutsche Forschungsgemeinschaft (DFG) im Rahmen der Exzellenzstrategie des Bundes und der Länder - EXC 2180 - 390900677 [HB], Department of Internal Medicine, Evangelische Kliniken Bonn gGmbH, Johanniter Krankenhaus, Bonn, Germany [Yon-Dschun Ko, Christian Baisch], Institute of Pathology, University of Bonn, Germany [Hans-Peter Fischer], Molecular Genetics of Breast Cancer, Deutsches Krebsforschungszentrum (DKFZ), Heidelberg, Germany [Ute Hamann], Institute for Prevention and Occupational Medicine of

the German Social Accident Insurance, Institute of the Ruhr University Bochum (IPA), Bochum, Germany [Thomas Brüning, Beate Pesch, Sylvia Rabstein, Anne Lotz]; and Institute of Occupational Medicine and Maritime Medicine, University Medical Center Hamburg-Eppendorf, Germany [Volker Harth]. GLACIER thanks Kelly Kohut, Patricia Gorman, Maria Troy. HABCS thanks Michael Bremer. HEBCS thanks Johanna Kiiski, Carl Blomqvist, Kristiina Aittomäki, Rainer Fagerholm, Kirsimari Aaltonen, Karl von Smitten, Irja Erkkilä. HKBCS thanks Hong Kong Sanatorium and Hospital, Dr Ellen Li Charitable Foundation, The Kerry Group Kuok Foundation, National Institute of Health 1R03CA130065 and the North California Cancer Center for support. HMBCS thanks Peter Hillemanns, Hans Christiansen and Johann H. Karstens. HUBCS thanks Shamil Gantsev. ICICLE thanks Kelly Kohut, Michele Caneppele, Maria Troy. KARMA and SASBAC thank the Swedish Medical Research Counsel. KBCP thanks Eija Myöhänen, Helena Kemiläinen. kConFab/AOCS wish to thank Heather Thorne, Eveline Niedermayr, all the kConFab research nurses and staff, the heads and staff of the Family Cancer Clinics, and the Clinical Follow Up Study (which has received funding from the NHMRC, the National Breast Cancer Foundation, Cancer Australia, and the National Institute of Health (USA)) for their contributions to this resource, and the many families who contribute to kConFab. LMBC thanks Gilian Peuteman, Thomas Van Brussel, EvyVanderheyden and Kathleen Corthouts. MABCS thanks Snezhana Smichkoska, Emilija Lazarova (University Clinic of Radiotherapy and Oncology), Katerina Kubelka-Sabit, Mitko Karadjozov (Adzibadem-Sistina Hospital), Andrej Arsovski and Liljana Stojanovska (Re-Medika Hospital)for their contributions and commitment to this study. MARIE thanks Petra Seibold, Dieter Flesch-Janys, Judith Heinz, Nadia Obi, Alina Vrieling, Sabine Behrens, Ursula Eilber, Muhabbet Celik, Til Olchers and Stefan Nickels. MBCSG (Milan Breast Cancer Study Group): Paolo, Peterlongo, Siranoush Manoukian, Bernard Peissel, Jacopo Azzollini, Dario Zimbalatti, Daniela Zaffaroni, Bernardo Bonanni, Irene Feroce, Mariarosaria Calvello, Aliana Guerrieri Gonzaga, Monica Marabelli, Davide Bondavalli and the personnel of the Cogentech Cancer Genetic Test Laboratory. The MCCS was made possible by the contribution of many people, including the original investigators, the teams that recruited the participants and

continue working on follow-up, and the many thousands of Melbourne residents who continue to participate in the study. We thank the coordinators, the research staff and especially the MMHS participants for their continued collaboration on research studies in breast cancer. MYBRCA thanks study participants and research staff (particularly Patsy Ng, Nurhidayu Hassan, Yoon Sook-Yee, Daphne Lee, Lee Sheau Yee, Phuah Sze Yee and Norhashimah Hassan) for their contributions and commitment to this study. The following are NBCS Collaborators: Anne-Lise Børresen-Dale (Prof. Em.), Kristine K. Sahlberg (PhD), Lars Ottestad (MD), Rolf Kåresen (Prof. Em.) Dr. Ellen Schlichting (MD), Marit Muri Holmen (MD), Toril Sauer (MD), Vilde Haakensen (MD), Olav Engebråten (MD), Bjørn Naume (MD), Alexander Fosså (MD), Cecile E. Kiserud (MD), Kristin V. Reinertsen (MD), Åslaug Helland (MD), Margit Riis (MD), Jürgen Geisler (MD), OSBREAC and Grethe I. Grenaker Alnæs (MSc). NBHS and SBCGS thank study participants and research staff for their contributions and commitment to the studies. OBCS thanks Arja Jukkola-Vuorinen, Mervi Grip, Saila Kauppila, Meeri Otsukka, Leena Keskitalo and Kari Mononen for their contributions to this study. OFBCR thanks Teresa Selander, Nayana Weerasooriya. ORIGO thanks E. Krol-Warmerdam, and J. Blom for patient accrual, administering questionnaires, and managing clinical information. The LUMC survival data were retrieved from the Leiden hospital-based cancer registry system (ONCDOC) with the help of Dr. J. Molenaar. PBCS thanks Louise Brinton, Mark Sherman, Neonila Szeszenia-Dabrowska, Beata Peplonska, Witold Zatonski, Pei Chao, Michael Stagner. The ethical approval for the POSH study is MREC /00/6/69, UKCRN ID: 1137. We thank staff in the Experimental Cancer Medicine Centre (ECMC) supported Faculty of Medicine Tissue Bank and the Faculty of Medicine DNA Banking resource. PREFACE thanks Sonja Oeser and Silke Landrith. PROCAS thanks NIHR for funding. RBCS thanks Jannet Blom, Saskia Pelders, Annette Heemskerk and the Erasmus MC Family Cancer Clinic. SBCS thanks Sue Higham, Helen Cramp, Dan Connley, Ian Brock, Sabapathy Balasubramanian and Malcolm W.R. Reed. We thank the SEARCH and EPIC teams. SGBCC thanks the participants and research coordinator Ms Tan Siew Li. SKKDKFZS thanks all study participants, clinicians, family doctors, researchers and technicians for their contributions and commitment to

this study. We thank the SUCCESS Study teams in Munich, Duessldorf, Erlangen and Ulm. SZBCS thanks Ewa Putresza. UCIBCS thanks Irene Masunaka. We thank the registration team of the Netherlands Comprehensive Cancer Organisation (IKNL) for the collection of data for the Netherlands Cancer Registry as well as IKNL staff for scientific advice.

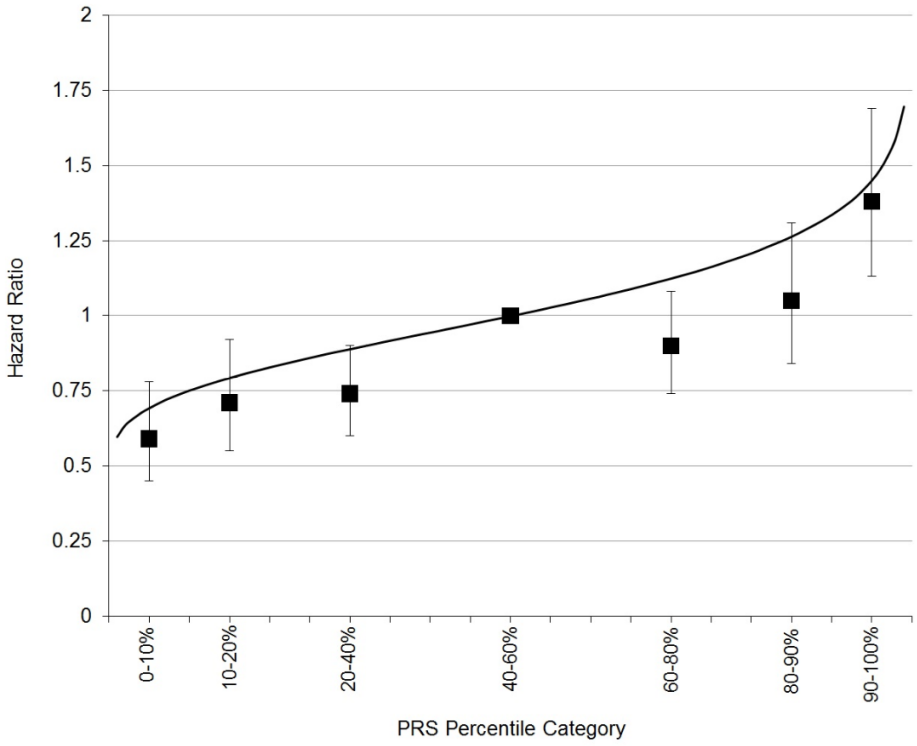
Supplemental References

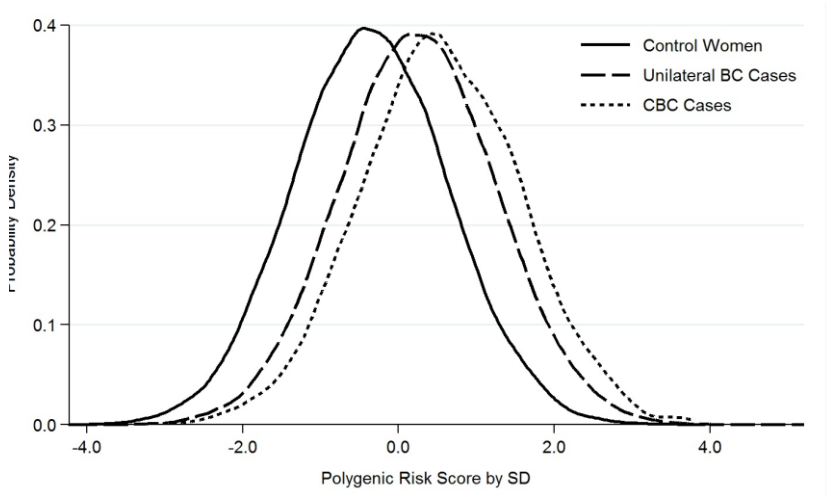
1. Mavaddat, N., Michailidou, K., Dennis, J., Lush, M., Fachal, L., Lee, A., Tyrer, J.P., Chen, T.H., Wang, Q., Bolla, M.K., et al. (2019). Polygenic Risk Scores for Prediction of Breast Cancer and Breast Cancer Subtypes. American journal of human genetics 104, 21-34.

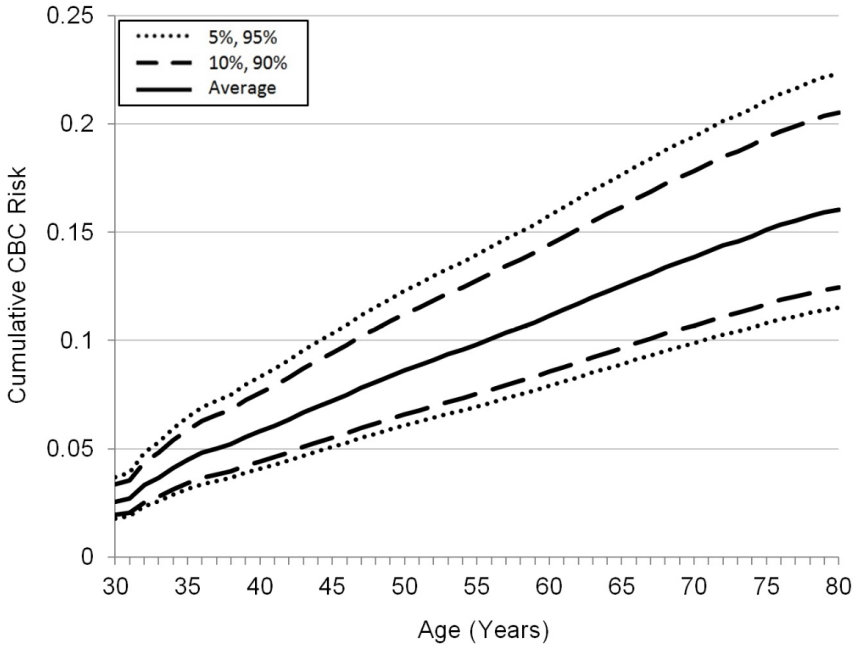
2. Kramer, I., Schaapveld, M., Oldenburg, H.S.A., Sonke, G.S., McCool, D., van Leeuwen, F.E., Van de Vijver, K.K., Russell, N.S., Linn, S.C., Siesling, S., et al. (2019). The influence of adjuvant systemic regimens on contralateral breast cancer risk and receptor subtype. J Natl Cancer Inst.

3. Mavaddat, N., Pharoah, P.D., Michailidou, K., Tyrer, J., Brook, M.N., Bolla, M.K., Wang, Q., Dennis, J., Dunning, A.M., Shah, M., et al. (2015). Prediction of breast cancer risk based on profiling with common genetic variants. J Natl Cancer Inst 107.

4. O'Connell, J., Sharp, K., Shrine, N., Wain, L., Hall, I., Tobin, M., Zagury, J.F., Delaneau, O., and Marchini, J. (2016). Haplotype estimation for Biobank-scale data sets. Nature genetics 48, 817-820.


Hazard Ratio
2
1.75
1.5
1.25
1
0.75
0.5
0.25
0
0-10%
10-20%
20-40%
40-60%
60-80%
80-90%
90-100%
PRS Percentile Category


Probability Density
0.4
0.3
0.2
0.1
0.0
-4.0
-2.0
0.0
2.0
4.0
Polygenic Risk Score by SD
Control Women
Unilateral BC Cases
CBC Cases


Cumulative CBC Risk
0.25
0.2
0.15
0.1
0.05
0
30
35
40
45
50
55
60
65
70
75
80
Age (Years)
5%, 95%
10%, 90%
Average

Study	Abbreviation	Country	Studydesign	Case definition	Control definition	How was follow-up (including vital status) information obtained?	When was the most recent attempt to have complete follow-up?	References
Australian Breast Cancer Family Study	ABCFS	Australia	Population-based case-control study	All cases diagnosed < age 40 plus a random sample of those diagnosed ages 40-59 from cancer registries in Victoria and New South Wales, plus a limited number diagnosed aged 60-69; cases living in Melbourne recruited from 1992-99 and in Sydney from 1993-98	Identified from the electoral rolls in Melbourne from 1992-98 and Sydney from 1993-99. Frequency matched to cases by age in 5 year categories	Systematic follow-ups by mail and telephone	2014	Dite, G.S. et al. Familial risks, early-onset breast cancer, and BRCA1 and BRCA2 germline mutations. J. Natl. Cancer Inst., 95, 448-57 (2003)
Amsterdam Breast Cancer Study	ABCS	Netherlands	Hospital-based consecutive cases; population-based controls (for iCOGS/OncoArray from blood bank)	Pre-iCOGS: All breast cancer patients (with operable, invasive mammacarcinoma) aged <50 years and diagnosed from 1970-1994 in four Dutch hospitals iCOGS/OncoArray: Breast cancer patients diagnosed before age 50 in 1995-2011 at the Netherlands Cancer Insitute - Antoni van Leeuwenhoek hospital (NKI-AVL)	Pre-iCOGS: Randomly selected women <50 years of age at baseline from 2 population-based prospective studies - the Monitoring Project on Cardiovascular Risk Factors (1987-1991) and the Monitoring Project on Chronic Disease Risk Factors (1993-1997). These studies were run by National Institute for Public Health and the Environment, The Netherlands. Controls are from the same catchment area as the breast cancer cases iCOGS/OncoArray: Population-based cohort of women recruited through the Sanquin blood bank, all ages	Hospital medical registry and linkage with municipality registry	January 2014	1) M. K. Schmidt, et al. Breast Cancer Survival and Tumor Characteristics in Premenopausal Women Carrying the CHEK2*1100delC Germline Mutation, J.Clin.Oncol. 25, 64-9 (2007). 2) K. Michailidou, et al. Large-scale genotyping identifies 41 new loci associated with breast cancer risk, Nat Genet. 45, 353-61 (2013)
Amsterdam Breast Cancer Study - Familial	ABCS-F	Netherlands	Clinical Genetic Center-based cases	Pre-iCOGS: Only in BCAC Phase I/II familial non-BRCA1/2 cases <50 from the Clinical Genetic Centre of the Netherlands Cancer Institute were included iCOGS/OncoArray: All non-BRCA1/2 breast cancer cases from the family cancer clinic of the NKI-AVL tested in the period 1995-2009; all ages and diagnosed with breast cancer in 1965-2012	No controls [Use controls of ABCS]	Hospital medical registry and linkage with municipality registry	April 2014	M.K. Schmidt, et al. Age- and Tumor Subtype-Specific Breast Cancer Risk Estimates for CHEK2*1100delC Carriers. J. Clin. Oncol. Jun 6. pii: JCO665844 (2016)
Australian Breast Cancer Tissue Bank	ABCTB	Australia	Hospital-based multi site newly diagnosed breast cancer case	Newly diagnosed unselected cases from 32 hospitals in New South Wales from 2006	Female controls from the Hunter Community Study (HCS) which is a population-based cohort study that consists of men and women aged 55-85 years of age who resided in Newcastle, New South Wales, Australia. Participants were randomly selected from the electroal role and contacted between Dec 2004 and Dec 2007. This study was conducted to assess factors important in the health, well-being, social functioning and economic consequences of ageing	Multiple (medical records at surgeron's rooms, medical records and databases at hospital clinics and GP Clincs)	Follow-up ongoing; Follow-up is attempted at 1 and 5 years after recruitment to the ABCTB, and 5 year intervals after this	McEvoy M, et al. Cohort profile: The Hunter Community Study. Int J Epidemiol. 2010 Dec;39(6):1452-63
Bavarian Breast Cancer Cases and Controls	BBCC	Germany	Hospital-based cases; population based controls	Consecutive, unselected cases with invasive breast cancer recruited at the University Breast Centre, Franconia in Northern Bavaria during 1999-2013	Healthy women with no diagnosis of cancer aged 55 or older. Invited by a newspaper advertisement in Northern Bavaria, and recruited during 1999-2013	Cancer registry and Medical records	12/1/2013	1) Fasching PA, et al. Single nucleotide polymorphisms of the aromatase gene (CYP19A1), HER2/neu status, and prognosis in breast cancer patients. Breast Cancer Res Treat. DOI 10.1007/s10549-007-9822-2 (2007). 2) Schrauder M, et al. Single nucleotide polymorphism D1853N of the ATM gene may alter the risk for breast cancer. J. Cancer Res. Clin. Oncol.,134, 873-82 (2008)
British Breast Cancer Study	BBCS	UK	Cancer registry and National Cancer Research network (NCRN) based cases; population-based controls	1) English & Scottish Cancer Registries: all breast cancer cases who developed a first primary before age 65 in 1971 or later and who subsequently developed a second primary cancer. 2) Unilateral breast cancer cases diagnosed before age 70 in 1971 or later	1) A friend, sister-in-law, daughter-in-law or other non-blood relative of cases. Recruitment of cases and controls began in January 2001	Cases are not followed up	We have not attempted to follow up	1) Johnson, N. et al. Interaction between CHEK2*1100delC and other low-penetrance breast-cancer susceptibility genes: a familial study. Lancet, 366, 1554-7 (2005). 2) Fletcher, O. et al. Inconsistent association between the STK15 F31I genetic polymorphism and breast cancer risk. J. Natl. Cancer. Inst., 98, 1014-8 (2006)
Breast Cancer Employment and Environment Study	BCEES	Australia	Population-based case-control study	First incident invasive breast cancer diagnosed between May 2009 and January 2011, residing in Western Australia and reported to the state wide mandatory Cancer Registry	Randomly selected from Western Australia electoral roll (registration is compulsory for Australian citizens)	No follow-up has been completed	No follow-up has been completed	Fritschi L, Erren TC, Glass DC, Girschik J, Thomson AK, Saunders C, et al. The association between different night shiftwork factors and breast cancer: a case-control study. Br J Cancer. 2013;109:2472-80
New York Breast Cancer Family Registry	BCFR-NY	USA	Clinic-based recruitment of families; family-based cohort	Recruitment took place from Jan 1996 to Dec 2012. Eligibility was based on one or more of the following criteria: two or more relatives with a personal history of breast or ovarian cancer; a woman diagnosed with breast or ovarian cancer at a young age; a woman with a history of both breast and ovarian cancer; an affected male; or known BRCA1 or BRCA2 mutation carriers	Unaffected family members also enrolled Jan 1996 to Dec 2012	Systematic follow-up every five years for questionnaire data and annual update of cancer history and vital records by at least one family member	Ongoing	1) John, EM et al. The Breast Cancer Family Registry: an infrastructure for cooperative multinational, interdisciplinary and translational studies of the genetic epidemiology of breast cancer. Breast Cancer Res 6, R375-R389 (2004). 2) Wu, HC et al. Repetitive element DNA methylation levels in white blood cell DNA from sisters discordant for breast cancer from the New York site of the Breast Cancer Family Registry. Carcinogenesis. 2012 Oct;33(10):1946-52. 3) Quante et al. Practical problems with clinical guidelines for breast cancer prevention based on remaining lifetime risk. J Natl Cancer Inst. 2015 May 8;107(7)
Philadelphia Breast Cancer Family Registry	BCFR-PA	USA	Clinic-based recruitment of families; family-based cohort	Recruitment took place from 1996 to 2011. Eligibility was based on one or more of the following criteria: 2 or more relatives with a personal history of breast or ovarian cancer; a woman diagnosed with breast or ovarian cancer at a young age; a woman with a history of both breast and ovarian cancer; an affected male; or known BRCA1 or BRCA1 mutation carriers	Unaffected family members also enrolled 1996 to 2011	Self-report on questionnaires or clinical database	2015 questionnaire	1) John, EM et al. The Breast Cancer Family Registry: an infrastructure for cooperative multinational, interdisciplinary and translational studies of the genetic epidemiology of breast cancer. Breast Cancer Res 6, R375-R389 (2004). 2) Terry M, Phillips K, Daly M, et al. Cohort Profile: The Breast Cancer Prospective Family Study Cohort (ProF-SC). International Journal of Epidemiology 2015;:dyv118
Utah Breast Cancer Family Registry	BCFR-UT	USA	Clinic-based recruitment of non-BRCA1/2 familial breast cancer cases;unaffected BRCA1/2 carriers as controls	Index cases from families tested negative for BRCA1/2 mutations. Recruited in Utah during 1995-2008	Unaffected BRCA1/2 carriers. Recruited in Utah during 1995-2008	Follow-up questionnaire	15 years after enrollment	1) John, EM et al. The Breast Cancer Family Registry: an infrastructure for cooperative multinational, interdisciplinary and translational studies of the genetic epidemiology of breast cancer. Breast Cancer Res 6, R375-R389 (2004). 2) Terry M, Phillips K, Daly M, et al. Cohort Profile: The Breast Cancer Prospective Family Study Cohort (ProF-SC). International Journal of Epidemiology 2015;:dyv118
Breast Cancer in Northern Israel Study	BCINIS	Israel	Population-based case-control study	All consecutive cases of invasive BC and DCIS diagnosed in a geographically defined area of Northern Israel since 1990. On-going	Age-sex-residence-ethnicity (Jews/Arabs)-matched controls, randomally sampled from the population using population registries. Recruitment since 1990 and on-going. A matched control will be interviewed within 6 months of the interview of its matched case	Oncology files, medical records, population database	June 2016	1) Rennert G, Pinchev M, Rennert HS. Use of Bisphosphonates and Risk of Postmenopausal Breast Cancer. J Clin Oncol 2010 Aug 1;28(22):3577-81. 2) Rennert G, Lejbkowicz F, Cohen I, Pinchev M, Rennert HS, Barnett-Griness O. MutYH mutation carriers have increased breast cancer risk. Cancer. 2011 Sep 22. doi: 10.1002/cncr.26506. [Epub ahead of print]
Breast Cancer in Galway Genetic Study	BIGGS	Ireland	Hospital-based cases; population based controls	Unselected cases recruited from West of Ireland since 2001. Cases were recruited from University College Hospital Galway and surrounding hospitals	Women > 60 years with no personal history of any cancer and no family History of breast or ovarian cancer were identified from retirement groups in the West of Ireland (same catchment area as cases) during the period 2001-2008	Local clinical database	01/08/2013	1) Colleran G, McInerney N, Rowan A, Barclay E, Jones AM, Curran C, Miller N, Kerin M, Tomlinson I, Sawyer E. The TGFBR1*6A/9A polymorphism is not associated with differential risk of breast cancer. Breast Cancer Res Treat. 2009 Apr 24. 2) Niall Mcinerney, Gabrielle Colleran, Andrew Rowan, Axel Walther, Ella Barclay, Sarah Spain Angela M. Jones Stephen Tuohy, Catherine Curran, Nicola Miller, Michael Kerin, Ian Tomlinson, Elinor J. Sawyer. Low penetrance breast cancer predisposition SNPs are site specific. Breast Cancer Research Treatment 2008 Nov 13 online
Breast Oncology Galicia Network	BREOGAN	Spain	Population-based case-control	A population-based study conducted since 1997 in two cities in Galicia, Spain (Vigo and Santiago) covering approximately 700,000 inhabitants. The study currently includes over 1600 incident breast cancer cases diagnosed from 1997-2014 in two Galician hospitals with blood, tumor tissue and risk factor questionnaire	Controls were frequency-matched to cases according to 5-year age group, inclusion in the universal Galician Public Health Service (SERGAS) registry database, and place of residence. They were healthy, unrelated female individuals from the same base population as cases randomly selected from SERGAS´ primary healthcare centers in the health areas of Santiago and Vigo. Recruitment began in 1997	From the computerised electronic individual medical history program unique for each individual in the state: Each individual is identified by his/her social security number which is unique to each person during his/her lifetime, his/her first name and his/her two last names, his/her national id number which is unique to each individual during his/her lifetime, date of bith, etc: All information is 100% computerised under the lanus program. This program includes information from all services and departments in all state hospitals, continuing update from all hospitals and family clinics as well as electronic prescriptions and medical imaging	Every six months	1) Castelao JE, Jiang XJ, Chavez-Uribe E, Fernandez Rodriguez B, Celeiro Muñoz C, Redondo Marey C, Peña Fernandez M, Novo Dominguez A, Dono Perez A, Doris Pereira C, Tomé MA, García-Caballero T, Martínez ME, Fraga Rodriguez M, Carracedo A, Forteza-Vila J, Gago-Dominguez M. Family History and Breast Cancer Subtypes in a Spanish Cohort PLoS One. 2012;7:e29459. 2) Redondo, C. M., M. Gago-Dominguez, et al. (2012). Breast feeding, parity and breast cancer subtypes in a Spanish cohort. PLoS One 7: e40543. 3) Ali, A. M., M. K. Schmidt, et al. (2014). Alcohol Consumption and Survival after a Breast Cancer Diagnosis: A Literature-Based Meta-analysis and Collaborative Analysis of Data for 29,239 Cases. Cancer Epidemiol Biomarkers Prev2014: 17. 4) Cruz, G. I., M. E. Martinez, et al. (2012). Hypothesized role of pregnancy hormones on HER2+ breast tumor development. Breast Cancer Res Treat: 8. 5) Gago-Dominguez et al. SpringerPlus (2016) 5:39 DOI 10.1186/s40064-015-1630-2 RESEARCH Alcohol and breast cancer tumor subtypes in a Spanish Cohort
Breast Cancer Study of the University of Heidelberg	BSUCH	Germany	Hospital-based cases;healthy blood donator controls	Cases diagnosed with breast cancer/breast cancer metastasis in 2008-2011 at the University Women´s Clinic Heidelberg	Healthy, unrelated, ethnically matched female blood donors recruited in 2007, 2009 & 2012 by German Red Cross Blood Service of Baden-Württemberg-Hessen, Institute of Transfusion Medicine & Immunology, Mannheim	Individual clinical investigation of patients, medical records	January 2017	Yang,R. et al. Genetic variants within miR-126 and miR-335 are not associated with breast cancer risk. Breast Cancer Res Treat 127, 549-554 (2011)
Canadian Breast Cancer Study	CBCS	Canada	Population-based case-control study	British Columbia (BC)- Incident cases diagnosed 2005-2009, resident in Vancouver area ascertained from the population cancer registry. Ontario - recruited from the Hotel Dieu Breast Assessment Program in Kingston, Ontario, 2005-2009	BC - cancer-free women who consented to participate in research studies through routine screening mammography (available to women in BC aged 40–79 years through the Screening Mammography Program of BC) in 2007-2008, resident in the same geographic areas and frequency-matched to cases by 5-year age group. Ontario - women with either normal mammogram results or a diagnosis of benign breast disease, frequency matched by 5-year age group recruited from the Hotel Dieu Breast Assessment Program in Kingston, Ontario, 2005-2009	Annual GP letter follow-up for patients discharged from BCCA (only for those who attended a BCCA clinic - ~85% of incident breast cancer cases); vital stats info primarily obtained through the British Columbia Vital Statistics Agency	Ongoing	1) Grundy A, Schuetz JM, Lai AS, Janoo-Gilani R, Leach S, Burstyn I, Richardson H, Brooks-Wilson A, Spinelli JJ, Aronson KJ. Shift work, circadian gene variants and risk of breast cancer. Cancer Epidemiol. 2013 Oct;37(5):606-12. doi: 10.1016/j.canep.2013.04.006. Epub 2013 May 28. PMID: 23725643. 2) Kobayashi LC, Janssen I, Richardson H, Lai AS, Spinelli JJ, Aronson KJ. Moderate-to-vigorous intensity physical activity across the life course and risk of pre- and post-menopausal breast cancer. Breast Cancer Res Treat. 2013 Jun;139(3):851-61. doi: 10.1007/s10549-013-2596-9. Epub 2013 Jun 15. PMID: 23771716. 3) 221. Grundy A, Richardson H, Burstyn I, Lohrisch C, Sengupta SK, Lai AS, Lee D, Spinelli JJ, Aronson KJ. Increased risk of breast cancer associated with long-term shift work in Canada. Occup Environ Med. (with commentary) 2013 Dec;70(12):831-8. doi: 10.1136/oemed-2013-101482. Epub 2013 Jul 1. PMID 23817841. 4) 227. Kobayashi LC, Janssen I, Richardson H, Lai AS, Spinelli JJ, Aronson KJ. A case-control study of lifetime light intensity physical activity and breast cancer risk. Cancer Causes Control 2014 Jan 25:133–140 DOI 10.1007/s10552-013-0312-z. Epub 2013 Oct 25. PMID: 24158779
Crete Cancer Genetics Program	CCGP	Greece	Hospital-based case-control study	Incident breast cancer cases treated between 2004 and 2013 at the University Hospital of Heraklion on Crete; all enrolled within 6 months of diagnosis	Healthy, unrelated, ethnically matched female blood donors recruited in 2014 by the laboratory of Hemostasis at the General Hospital of Heraklion "Venizelio"	Individual patient medical records	August 2014 for all patietns included in the database	No references yet
Copenhagen General Population Study	CGPS	Denmark	Population-based case-control study	Consecutive, incident cases from 1 hospital with centralized care for a population of 400,000 women from 2001 to the present	Community controls residing in the same region as cases and with no history of breast cancer were identified from the Copenhagen General Population Study recruited 2003-2007. All controls were known to still be breast cancer-free at the end of 2007	Vital status: from the citizen registry	01/12/2017	Weischer,M., Bojesen,S.E., Tybjaerg-Hansen,A., Axelsson,C.K., & Nordestgaard,B.G. Increased risk of breast cancer associated with CHEK2*1100delC. J Clin Oncol 25, 57-63 (2007)
Spanish National Cancer Centre Breast Cancer Study	CNIO-BCS	Spain	Case-control study	Two groups of cases:1) 574 consecutive breast cancer patients, unselected for family history, from 3 public hospitals, 2 in Madrid and one in Oviedo, from 2000 to 2005. 2) 291 cases with at least one first degree relative also affected with breast cancer, recruited through the CNIO family cancer clinic in Madrid from 2000 to 2004	Women attending the Menopause Research Centre between 2000 and 2004 and female members of the College of Lawyers attending a free, targeted medical check-up in 2005, all free of breast cancer and all in Madrid	Only available for a subset of cases (using clinical records?)	2007/2008	Milne, RL et al. ERCC4 associated with breast cancer risk: a two stage case-control study using high throughput genotyping. Cancer Res., 66, 9420-7 (2006)
California Teachers Study	CTS*	USA	Prospective cohort study: nested case-control	This is a nested case-control study conducted within a cohort of California teachers(113,590) who were under age 80 years at baseline, had no prior history of invasive or in situ breast cancer. Cases are women newly diagnosed with a histologically confirmed invasive primary adenocarcinoma of the breast at age 80 years or younger from 1998 to 2008	Controls are a probability sample of at-risk cohort members, frequency matched to cases on age at baseline (5-year age groups), self-reported race/ethnicity (white, African American, Latina, Asian, other), and broad geographic region within California. Controls were recruited during 1998 to 2008 and selected without replacement, using an assigned reference date	The vital status and follow-up date are standard items of the California Cancer Registry	Completed information is as of 10/30/2014	Bernstein L, Allen M, Anton-Culver H, Deapen D, Horn-Ross PL, Peel D, Pinder R, Reynolds P, Sullivan-Halley J, West D et al. High breast cancer incidence rates among California teachers: results from the California Teachers Study (United States). Cancer Causes Control 2002, 13(7):625-635
DietCompLyf Breast Cancer Survival Study	DIETCOMPLYF	UK	Multi-centre prospective cohort study	Invasive primary breast cancer grade I-III, patients recruited 9 - 15 months after diagnosis, <age 75. Recruitment throughout UK. Patient first recruited on 18/2/97. Study joined NCRN in July 2004. Recruitment finished on 31/8/10	No controls	From the patients at the last time they visited the hospital	28th March 2014	Swann R, Perkins KA, Velentzis LS, Cria C, Dutton SJ, Mulligan AA, Woodside JV, Cantwell MM, Leathem AJ, Robertson CE, Dwek MV. The DietCompLyf study: A prospective cohort study of breast cancer survival and phytoestrogen consumption. Maturitas (2013) 75: 232-240
Family History Risk Study	FHRISK	UK	Clinic-based cohort study with a nested case-control study	Women diagnosed with breast cancer and attending the Family History Clinic in Manchester for increased risk of breast cancer. Recruitment period 2009-2012	Women attending the same Family History Clinic as the cases but without a breast cancer diagnosis. Recruitment period is the same as for the cases	Follow up continues until the participant has been discharged from the Family History Clinic	Ongoing	1) Evans DG, Astley S, Stavrinos P, Harkness E, Donnelly LS, Dawe S, Jacob I, Harvie M, Cuzick J, Brentnall A, Wilson M, Harrison F, Payne K, Howell A. Improvement in risk prediction, early detection and prevention of breast cancer in the NHS Breast Screening Programme and family history clinics: a dual cohort study. Southampton (UK): NIHR Journals Library; 2016 Aug. 2) Ingham SL, Warwick J, Buchan I, Sahin S, O'Hara C, Moran A, Howell A Evans DG. Ovarian cancer among 8005 women from a breast cancer family history clinic: No increased risk of invasive ovarian cancer in families testing negative for BRCA1 and BRCA2. J Med Genet 2013;50(6):368-72

German Consortium for Hereditary Breast & Ovarian Cancer	GC-HBOC*	Germany	Clinic-based case study and prospective cohort study	Women diagnosed with breast cancer in one of the GC-HBOC centres (Cologne, Munich, Kiel, Heidelberg, Düsseldorf, Ulm, Würzburg, Münster and Hannover). Recruitment period 1996-present	Healthy, unrelated, ethnically and age-matched female control individuals (LIFE study, Leipzig, Germany)	Medical records or personal visit for women under intensified surveillance	Updated at least annually for women under intensified surveillance	1) Kast K, et al. Prevalence of BRCA1/2 germline mutations in 21,401 families with breast and ovarian cancer. J Med Genet 2016 53(7):465-71. 2) Rhiem K, et al. Breast Cancer Res 2012 Dec 7;14(6):R156. 3) Graeser MK, et al. J Clin Oncol 2009 Dec 10;27(35):5887-92. 4) Engel C, et al. BMC Cancer 2018 Mar 7;18(1):265
Gene Environment Interaction and Breast Cancer in Germany	GENICA*	Germany	Population-based case-control study	Incident breast cancer cases enrolled between 2000 and 2004 from the Greater Bonn area (by of the hospitals within the study region); all enrolled within 6 months of diagnosis	Selected from population registries from 31 communities in the greater Bonn area; matched to cases in 5-year age classes between 2001 and 2004	Through telephone interview with the patient or patient's relative, as well as information from the registration office and clinical records	Year 2012	1) Pesch, B. et al. Factors modifying the association between hormone –replacement therapy and breast cancer risk. Eur. J. Epidemiol., 20:699-711 (2005). 2) Justenhoven,C. et al. The CYP1B1_1358_GG genotype is associated with estrogen receptor-negative breast cancer. Breast Cancer Res Treat. 111, 171-177 (2008)
Genetic Epidemiology Study of Breast Cancer by Age 50	GESBC	Germany	Population-based study of women <50 years	All incident cases diagnosed <50 years of age in 1992-5 in two regions: Rhein-Neckar-Odenwald and Freiburg, by surveying the 38 clinics serving these regions	Selected from random lists of residents of the study regions supplied by population registries; two controls were selected for each case, matched by age and study region. Recruitment was carried out 1992-1998	Vital status was obtained by requesting this information from the population registry	2009, but only for vital status and cause of death	Chang-Claude,J., Eby,N., Kiechle,M., Bastert,G., & Becher,H. Breastfeeding and breast cancer risk by age 50 among women in Germany. Cancer Causes Control 11, 687-695 (2000)
Study to Investigate the Genetics of Lobular Carcinoma In situ in Europe	GLACIER	UK	Hospital-based case-control study	Cases aged 60 or younger with LCIS (pure or associated with invasive cancer of any subtype) or invasive lobular breast cancer from 96 hospitals throughout the UK. Recruitment period was from Jun 2007 to Sep 2012	Controls were healthy women of any age with no history of LCIS, DCIS, breast disease or invasive breast cancer and who had no close relative (mother, sister, daughter or aunt) who had been so affected. Controls were recruited by asking non-blood relatives (generally sisters-in-law or friends) of affected individuals to act as a control. Recruitment period was from Jun 2007 to Jun 2013	Self reported by patient at time of recruitment	Not reported	Sawyer E et al. Genetic predisposition to in situ and invasive lobular carcinoma of the breast. PLoS Genet. 2014 Apr 17;10(4)
Hannover Breast Cancer Study	HABCS	Germany	Hospital-based case-control study	Cases who received radiotherapy for breast cancer at Hannover Medical School between 1996-2003 (HaBCS I), or were diagnosed with breast cancer at a certified Breast Cancer Clinics in the Hannover region between 2012-2016 (HaBCS II), unselected for age or family history	Anonymous female blood bank donors at Hannover Medical School, collected from 8/2005-12/2005, with known age and ethnic background	Follow-up information was obtained through the central tumour registry at MHH; rarely through telephone contact with clinicians	Summer 2016	Dork, T. et al. Spectrum of ATM gene mutations in a hospital-based series of unselected breast cancer patients. Cancer Res., 61, 7608-7615 (2001)
Hospital Clínico San Carlos	HCSC	Spain	Population-based study of priori sporadic breast cancer cases	This is a cohort of a priori sporadic breast cancer patients which includes a cohort of 200 patients that were enroled in a neoadyuvant trial in which women were randomized to neoadyuvant docetaxel vs neoadyuvant doxorrubicin. Recruitment period was from 2000 to 2013. Most patients have been treated in the Clinio San Carlos hospital (Madrid) and most and are being followed in different hospitals of Madrid	No controls	Medical records	Feburary 2018	1) Romero A, García-Saenz JA, Fuentes-Ferrer M, López García-Asenjo JA, Furió V, Román JM, Moreno A, de la Hoya M, Díaz-Rubio E, Martín M, Caldés T. Correlation between response to neoadjuvant chemotherapy and survival in locally advanced breast cancer patients. Ann Oncol. 2013 Mar;24(3):655-61. 2) Martín M, Romero A, Cheang MC, López García-Asenjo JA, García-Saenz JA, Oliva B, Román JM, He X, Casado A, de la Torre J, Furio V, Puente J, Caldés T, Vidart JA, Lopez-Tamuella S, Diaz-Rubio E, Perou CM. Genomic predictors of response to doxorubicin versus docetaxel in primary breast cancer. Breast Cancer Res Treat. 2011 Jul;128(1):127-3
Helsinki Breast Cancer Study	HEBCS	Finland	Hospital-based case-control study, plus additional familial cases	(1) Consecutive cases (883) from the Department of Oncology, Helsinki University Central Hospital 1997-8 and 2000, (2) Consecutive cases (986) from the Department of Surgery, Helsinki Universily Central Hospital 2001 – 2004, (3) Familial breast cancer patients (536) from the Helsinki University Central Hospital, Departments of Oncology and Clinical Genetics (1995-)	Healthy females from the same geographical region in Southern Finland in 2003	Hospital medical records, Cancer registry, population registry	2011 breast cancer spec/ 2015 overall	1) Syrjakoski, K. et al. Population-based study of BRCA1 and BRCA2 mutations in 1035 unselected Finnish breast cancer patients. J. Natl. Cancer Inst., 92, 1529-31 (2000). 2) Kilpivaara, O. et al. Correlation of CHEK2 protein expression and c.1100delC mutation status with tumor characteristics among unselected breast cancer patients. Int J. Cancer, 113, 575-80 (2005). 3) Fagerholm,R. et al. NAD(P)H:quinone oxidoreductase 1 NQO1*2 genotype (P187S) is a strong prognostic and predictive factor in breast cancer. Nat Genet 40, 844-853 (2008)
Hospital-based Epidemiologic Research Program at Aichi Cancer Center	HERPACC	Japan	Hospital-based case-control study	Incident breast cancer cases who firstly visited Aichi Cancer Center between 2001 and 2013 and were diagnosed within 1 year from the first visit. No previous history of any type of cancer	Controls were selected from pool of non-cancer patients who firstly visited Aichi Cancer Center between 2001-2011. Non-cancer status is defined as "having no positive finding on any of clinical/laboratory/graphical examination within 1 year from their first visit. No previous history of cancer is allowed	Checking of medical records	End of 2014	Kawase T et al. FGFR2 intronic polymorphisms interact with reproductive risk factors of breast cancer: results of a case control study in Japan. Int J Cancer 2009: 125:1946-1952
Hong Kong Breast Cancer Study	HKBCS	Hong Kong	Hospital-based case-control study	Genetic screening of high risk breast cancer patients from all Hong Kong hospitals. Incidence cases clasified as high risk group: 1) first degree relative with breast and/or ovarian cancer, 2) cases where age is less than or equal to 45 years, 3) bilateral breast cancer, 4) triple negative breast cancer cases, 5) family history of breast and /or ovarian cancer. Cases were recruited 2006-2014	Controls were selected from pool of non-cancer patients who visited Hong Kong hospitals. Same period of recruitment as cases	Not reported	Not reported	1) Kwong A et al. Novel BRCA1 and BRCA2 genomic rearrangements in Southern Chinese breast/ovarian cancer patients. Breast Cancer Res Treat. 2012: 136(3):931-3. 2) Kwong A et al. Identification of BRCA1/2 founder mutations in Southern Chinese breast cancer patients using gene sequencing and high resolution DNA melting analysis. PLoS One. 2012: 7(9):e43994
Hannover-Minsk Breast Cancer Study	HMBCS	Belarus	Hospital-based cases; population based controls	Ascertainment at the Byelorussian Institute for Oncology and Medical Radiology Aleksandrov N.N. in Minsk or at one of 5 regional oncology centers in Gomel, Mogilev, Grodno, Brest or Vitebsk through the years 2002-2008	Controls from the same population aged 18-72 years. Healthy (without personally history of cancer) female probunds recruited from the same geographical regions as cases during the years 2002-2008. About 75% of controls were women invited for general medical examination at five regional gynecology clinics (in Gomel, Mogilev, Grodno, Brest or Vitebsk) and cancer-free volunteers ascertained at the Institute for Inherited Diseases in Minsk; 20% were cancer-free female blood bank donors recruited at Republic Blood Bank, Minsk, Belarus; finally 5% of controls were healthy cancer-free relatives of some breast cancer patients	No data provided	No data provided	Bogdanova,N. et al. A nonsense mutation (E1978X) in the ATM gene is associated with breast cancer. Breast Cancer Res Treat 118, 207-211 (2009)
Hannover-Ufa Breast Cancer Study	HUBCS	Russia	Hospital-based cases; population based controls	Consecutive Russian breast cancer patients aged 24-86 years ascertained at one of the two participating oncological centers in Bashkorstostan and Siberia through the years 2000-2008	Population controls aged 18-84 years recruited from a population study of different populations of Russia. Healthy volunteers (without any malignancy) were selected from the same geographical regions during the years 2002-2008	Medical records	Varies by participant	Bogdanova,N. et al. A nonsense mutation (E1978X) in the ATM gene is associated with breast cancer. Breast Cancer Res Treat.(2008)
Study to Investigate the Genetics of In Situ Carcinoma of the Ductal Subtype	ICICLE	UK	Hospital-based case-control study	Cases aged 60 or younger with pure DCIS (no associated invasive cancer of any subtype) from 96 hospitals throughout the UK. Recruitment period was from Jul 2008 to Nov 2012	Controls were healthy women of any age with no history of LCIS, DCIS, breast disease or invasive breast cancer and who had no close relative (mother, sister, daughter or aunt) who had been so affected. Controls were recruited by asking non-blood relatives (generally sisters-in-law or friends) of affected individuals to act as a control. Recruitment period was from Jun 2007 to Jun 2013	Self reported by patient at time of recruitment	Not reported	No references yet
Karolinska Breast Cancer Study	KARBAC	Sweden	Population and hospital-based cases; geographically matched controls	1. Familial cases from Department of Clinical Genetics, Karolinska University Hospital , Stockholm. 2. Consecutive cases from Department of Oncology, Huddinge & Söder Hospital, Stockholm 1998-2000	Blood donors of mixed gender from same geographical region. Excess material was received from all blood donors over a 3 month period in 2004 (approximately 3000) and DNA was extracted from a random sample of 1500	Medical records	2016	1) Wendt C. et al. Tumor spectrum in non-BRCA hereditary breast cancer families in Sweden. Hered Cancer Clin Pract. 2015 16;13(1):15. 2) Margolin S. et al. BRCA1 mutations in a population –based study of breast cancer in Stockholm County. Genet. Test., 8, 127-32 (2004)
Karolinska Mammography Project for Risk Prediction of Breast Cancer - Cohort Study	KARMA	Sweden	Cohort study	Inclusion of 70,877 women Oct 2010 - March 2013. 3000 women had BC at cohort entry. In all, 800 women have been diagnosed with breast cancer since study entry (Oct 2015). Approximately 250 women are diagnosed with BC annually	Non - BC cases in the Karma Cohort	Through the Swedish Cause of Death register, Clinical Breast Cancer register and the Inpatient register	We match the data sets to all registers twice a year	Submitted
Kuopio Breast Cancer Project	KBCP	Finland	Population-based prospective clinical cohort	1. Women seen at Kuopio University Hospital between 1990 and 1995 because of breast lump, mammographic abnormality, or other breast symptom who were found to have breast cancer. 2. Consecutive malignant breast cancer cases diagnosed at KUH from 2011 onwards	Age and long-term area-of-residence matched controls selected from the National Population Register and interviewed in parallel with the cases	Follow-up is done by an oncologist	2016	1) Hartikainen, J.M. et al. An autosome-wide scan for linkage disequilibrium-based association in sporadic breast cancer cases in eastern Finland: three candidate regions found. Cancer Epidemiol. Biomarkers Prev., 14, 75-80 (2005). 2) Hartikainen,J.M. et al. Refinement of the 22q12-q13 breast cancer-associated region: evidence of TMPRSS6 as a candidate gene in an eastern Finnish population. Clin Cancer Res 12, 1454-1462 (2006)
Kathleen Cuningham Foundation Consortium for research into Familial Breast Cancer/Australian Ovarian Cancer Study	kConFab/AOCS	Australia and New Zealand	Clinic-based recruitment of familial breast cancer patients (cases); population-based case-control study of ovarian cancer (controls only)	Cases were from multiple-case breast and breast-ovarian families recruited though family cancer clinics from across Australia and New Zealand from 1998 to the present. Cases were selected for inclusion in BCAC studies if (i) family was negative for mutations in BRCA1 and BRCA2 (ii) case was the index for the family, defined as youngest breast cancer affected family member	Female controls were ascertained by the Australian Ovarian Cancer Study identified from the electoral rolls from all over Australia from 2002-2006	Patient self and family reports, medical records	Ongoing, varies by participant	1) Mann, G.J. et al. Analysis of cancer risk and BRCA1 and BRCA2 mutation prevalence in the kConFab familial breast cancer resource. Breast Cancer Res., 8, R12 (2006). 2) Beesley, J et al. Association between single nucleotide polymorphisms in hormone metabolism and DNA repair genes and epithelial ovarian cancer: Results from two Australian studies and an additional validation set. Cancer Epidemiol. Biomarkers Prev., 12, 2257-65 (2007)
Leuven Multidisciplinary Breast Centre	LMBC	Belgium	Hospital-based case-control study	All patients diagnosed with breast cancer and seen in the Multidisciplinary Breast Center in Leuven (Gashuisberg) since June 2007 plus retrospective collection of cases diagnosed since 2000	Healthy controls (blood donors) collected at the Red Cross and located in Gasthuisberg hospital (Oct-2007-March 2008)	KWS; the latest data covers all departments, not only when they come for the breast pathology	2013 in the datasheet from BCAC. Normally it's in the KWS when the patient has been here	1) Neven P, Brouckaert O, Van Belle V, Vanden Bempt I, Hendrickx W, Cho H, Deraedt K, Van Calster B, Van Huffel S, Moerman P, Amant F, Leunen K, Smeets A, Wildiers H, Paridaens R, Vergote I, Christiaens MR. In early-stage breast cancer, the estrogen receptor interacts with correlation between human epidermal growth factor receptor 2 status and age at diagnosis, tumor grade, and lymph node involvement. J Clin Oncol. 2008 Apr 1;26(10):1769-71. 2) De Maeyer L, Van Limbergen E, De Nys K, Moerman P, Pochet N, Hendrickx W, Wildiers H, Paridaens R, Smeets A, Christiaens MR, Vergote I, Leunen K, Amant F, Neven P. Does estrogen receptor negative/progesterone receptor positive breast carcinoma exist? J Clin Oncol. 2008 Jan 10;26(2): 336-8
Macedonian Breast Cancer Study	MABCS	Republic of North Macedonia	Hospital-based case-control study	Prospectively ascertained cases of breast cancer in two Hospitals in Skopje, Macedonia from 2012 to 2014. Ethnic origin: Macedonians (~82.8%) and Albanians (~17.2%). Age of the cases: 29 to 86, mean 53.8	Pregnant women without breast cancer undergoing prenatal screening for chromosomal aneuploidy from 2013-2014. Recruited in three hospitals in Skopje, two of which are the same as those for recruitment of cases. Controls were matched for ethnic origin with the cases. Age of the controls: 18-45, mean 31.1	No information available	Not reported	No References
Mammary Carcinoma Risk Factor Investigation	MARIE	Germany	Population-based case-control study	Incident cases diagnosed from 2001-2005 in the study region Hamburg in Northern Germany, and from 2002-2005 in the study region Rhein-Neckar-Karlsruhe in Southern Germany	2 controls per case were randomly drawn from population registries and frequency matched by birth year and study region to the case. Controls were recruited from 2002 to 2006	Follow-up information was obtained through follow-up interviews/questionnaires and new events through medical records to verify clinical events either reported by treating physicians or self-reported during follow-up interviews. Vital status was obtained by requesting this information from the population registry	May 2016	Flesch-Janys, D et al.Risk of different histological types of postmenopausal breast cancer by type and regimen of menopausal hormone therapy Int J Cancer. 2008 Aug 15;123(4):933-41
Milan Breast Cancer Study Group	MBCSG	Italy	Clinic-based recruitment of familial/early onset breast cancer patients (cases); population-based controls	Familial and/or early onset breast cancer patients (aged 22-87) negative for mutations in BRCA genes, ascertained in two large cancer centres in Milan from 1996 to 2008	Healthy blood donors aged 18-71 years, retruited at two blood centres in Milan from 2004 (centre 1) and 2007 (centre 2) to 2009	80% Medical records; 5% Phone contact; 15% Referred by patients/family members	Informations on follow-up not routinely collected	1) De Vecchi et al. Evidences for association of the CASP8 -652 6N del promoter polymorphism with age at diagnosis in familial breast cancer cases (letter). Breast Cancer Res Treat 113:607-8, 2009. 2) Catucci et al. Letter to the editor: SNPs in ultraconserved elements and familial breast cancer risk. Carcinogenesis 30:544–545, 2009
Mayo Clinic Breast Cancer Study	MCBCS*	USA	Hospital-based case-control study	Incident cases residing in 6 states (MN, WI, IA, IL, ND, SD) seen at the Mayo Clinic in Rochester, MN from 2002-5	Women without cancer presenting for general medical examination at the Mayo Clinic. Controls were recruited concurrently with cases and were frequency matched to cases on age, ethnicity and county/state	Not reported	Not reported	Olson, JE. et al. A comprehensive examination of CYP19 variation and breast density. Cancer Epidemiol. Biomarkers Prev. 16, 623-5 (2007)
Melbourne Collaborative Cohort Study	MCCS	Australia	Prospective cohort study: nested case-control study	Incident cases diagnosed between baseline (1990-1994) and last follow-up (2012) among the 24469 women participating in the cohort	For each case a control was randomly selected from women from the cohort who did not develop breast cancer before the age at diagnosis of the case and matched the case on year of birth and country of birth	Record linkage to the national and state cancer and death registries	Record linkages are carried out at least annually	Giles GG. et al. The Melbourne Collaborative Cohort Study. IARC Sci. Publ., 156, 69-70 (2002)
Multiethnic Cohort	MEC	USA	Prospective cohort study: nested case-control	Incident cases identified from SEER cancer registries in Los Angeles County & State registries in California & Hawaii, USA from 1993-2002. Grouped by self-reported ethnicity	Women without cancer from the same States, recruited concurrently with cases & frequency matched to cases by age at blood-draw & self-reported ethnicity	Linkage to SEER registries, state vital statistics and National Death Index	Linkages are performed annually	Kolonel, L. N. et al. A multi-ethnic cohort in Hawaii and Los Angeles; Baseline characteristics. Am. J. Epidemiol., 151, 346-357 (2000)
Melanoma Inquiry of Southern Sweden	MISS	Sweden	Population-based prospective cohort study	Population based cohort off women aged 25-65 in southern Sweden, born in Sweden, no cancer diagnosis before, interviewed about cancer risk factors 1990,2000,2010, saliva sampled 2011, cancer incidence/mortality followed through registries	2 matched controls within the cohort	Cause of death registry, records	2016	1) Olsson HL, Ingvar C, Bladstrom A. Hormone replacement therapy containing progestins and given continuously increases breast carcinoma risk in Sweden. Cancer. 2003 Mar;97(6):1387-92. 2) Nielsen K, Másbäck A, Olsson H, Ingvar C. A prospective study of 40000 women regarding host factors, UV exposure and sunbed use in relation to risk and anatomic site of malignant melanoma Int J Cancer. 2012;131(3)Aug 1):706-15
Mayo Mammography Health Study	MMHS	USA	Prospective Cohort Study (2003-2006) of women ages 35+ receiving screening mammography at Mayo Clinic and living in MN, IA, WI; nested case-control	Incident cases (Invasive or in situ) diagnosed at least 3 months after enrollment	Two sets of controls. One set frequency matched to cases on age. Second set of premenopausal women with density measures	Multiple sources: linkage to registration/mailings	2014	Olson JE, Sellers TA, Scott CG, Schueler BA, Brandt KR, Serie DJ, Jensen MR, Wu FF, Morton MJ, Heine JJ, Couch FJ, Pankratz VS, Vachon CM. The influence of mammogram acquisition on the mammographic density and breast cancer association in the Mayo Mammography Health Study cohort. Breast Cancer Res. 2012 Nov 15;14(6):R147
Malaysian Breast Cancer Genetic Study	MYBRCA	Malaysia	Hospital-based case-control study	Breast cancer cases identified at the Breast Cancer Clinic in University Malaya Medical Centre Jan 2003-July 2014 and Subang Jaya Medical Centre Sep 2012-Sept 2014; cases are a mixture of prevalent and incident cases	Controls are cancer-free individuals (37-74 years) selected from women attending mammographic screening at the same hospitals	From National Registry of Births and Deaths	Annual exercise, in January or February	Tan M-M, Ho W-K, Yoon S-Y, Mariapun S, Hasan SN, Lee DS-C, et al. (2018) A case-control study of breast cancer risk factors in 7,663 women in Malaysia. PLoS ONE 13(9): e0203469. https://doi.org/10.1371/journal.pone.0203469

Norwegian Breast Cancer Study	NBCS*	Norway	Hospital-based case-control study	Incidence cases from three different hospitals: 1) Cases (114) mean age 64 (28-92) at Ullevål Univ. Hospital 1990-94, 2) cases (182) mean age 59 (26-75) referred to Norwegian Radium Hospital 1975-1986, 3) cases (124), mean age 56 (29-82) with stage I or II disease, in the Oslo micro-metastases study at Norwegian Radium Hospital between 1995-1998, 4) Breast cancer cases referred to the Norwegian hospitals Akershus University Hospital in Lørenskog, Ullevaal university hospital in Oslo and Rikshospitalet-Radiumhospitalet in Oslo from 2007-2010. Mean age is 63 years. Consecutive series. 5) Breast cancer cases referred to the Norwegian Radium Hospital hospitalet 2010-2013. Neoadjuvantly treated with Avastin (Bevacizumab). 6) Consecutive series of Breast cancer incidents referred to Akershus university hospital 2004-2014.	Control subjects were healthy women, age 55-71, residing in Tromsø (440), and Bergen (109) attending the Norwegian Breast Cancer Screening Program. Healthy tissue from mammoplastic reduction surgery at a private clinic in Oslo	Medical records	Every 5 years, follow up 5-20 years, different for all sub-cohorts	1) Aure et al. Genome Med. 2015 Feb 2;7(1):21. 2) Fleischer et al. 2014 Genome Biol. 2014;15(8):435. 3) Fleischer et al. 2014 Int J Cancer. 2014 Jun 1;134(11):2615-25. 4) Quigley et al. 2014 Mol Oncol. 2014 Mar;8(2):273-84
Nashville Breast Health Study	NBHS*	USA	Population-based case-control study	Through a rapid case-ascertainment system, we identified newly-diagnosed breast cancer cases through the Tennessee State Cancer Registry and five major hospitals in the city that provide medical care for breast cancer patients. Eligible cases were women diagnosed with invasive breast cancer or ductal carcinoma in situ, who were between the ages of 25 and 75, had no prior history of cancer other than non-melanoma skin cancer, had a resident telephone, spoke English, and who were able to provide consent to the study. Recruitment period was from 2001 to 2011. The recruitment for European Americans ended in 2008.	Controls were identified via random digit dialing (RDD) of households in the same geographic area as cases during 2001-2011. Eligibility criteria for controls were the same as cases with the exception that controls did not have a prior cancer diagnosis other than simple skin cancer. Controls were frequency matched to cases on 5-year age group, race, and county of residence	Not reported	Not reported	Zheng W, Long J, Gao YT, Li C, Zheng Y, Xiang YB, Wen W, Levy S, Deming SL, Haines JL, Gu K, Fair AM, Cai Q, Lu W, Shu XO. Genome-wide association study identifies a new breast cancer susceptibility locus at 6q25.1. Nature Genetics 41(3):324-8, 2009. PMCI
Northern California Breast Cancer Family Registry	NC-BCFR	USA	Population-based recruitment of families; family-based cohort; population-based controls for subset of cases	Incident breast cancer cases included women aged <65 years diagnosed from 1995-2009, identified through the SEER cancer registry of the Greater San Francisco Bay Area. All cases with indicators of increased genetic risk were eligible to enroll (dx at age <35 yrs, personal history of ovarian or childhood cancer, bilateral breast cancer with 1st dx at age <50, family history of breast or ovarian cancer in first-degree relatives). Cases not meeting these criteria were randomly sampled (2.5% of non-Hispanic whites, 32% of other race/ethnicities). Incident cases also included men aged <80 years diagnosed from 1995-1998	1) Unaffected family members enrolled from 1995-2011. 2) Unaffected unrelated population controls identified through random digit dialing conducted from 1999-2000 in the San Francisco Bay Area. Controls were frequency matched to cases diagnosed from 1995-1998 on 5-year age group and race/ethnicity, at a ratio of 1 control per 2 cases	Active follow-up by questionnaire and linkage with the California Cancer Registry. Annual phone follow-up from 1999-2012 to obtain updates on vital status and new cancers in the family. Updates on risk factors, vital status, and new cancers in family in 2007-2011, 2012-2014, and 2015-2017	2015-2017	1) John,E.M. et al. The Breast Cancer Family Registry: an infrastructure for cooperative multinational, interdisciplinary and translational studies of the genetic epidemiology of breast cancer. Breast Cancer Res 6, R375-R389 (2004). 2) Terry M, Phillips K, Daly M, et al. Cohort Profile: The Breast Cancer Prospective Family Study Cohort (ProF-SC). International Journal of Epidemiology 2015;:dyv118
North Carolina Breast Cancer Study	NCBCS	USA	NCBCS Phases 1 & 2: population-based case-control study NCBCS Phase 3: population-based case-only study In each phase, African American women and women under the age of 50 with invasive breast cancer were over-sampled. There was no over sampling by race or age for CIS cases in Phase 2	NCBCS Phase 1: women aged of 20-74 residing in the 24 North Carolina county area and diagnosed with a first primary invasive breast cancer from 1993-1996 NCBCS Phase 2: women aged of 20-74 residing in the same study area and diagnosed with a first primary invasive breast cancer from 1996-2000. All women diagnosed with DCIS, DCIS with microinvasion to a depth of 2mm, LCIS, and mixed of DCIS & LCIS from 1996-2000 were also eligible NCBCS Phase 3: Study area was expanded to 44 NC counties. Women aged 20-74 residing in this area and diagnosed with a first primary invasive breast cancer from May, 2008 to July, 2013	NCBCS Phase 1 & 2: Invasive: controls were from the same county area and frequency matched to cases by race (African American vs. non-AA) and five-year age-group (20-24, 25-29,.., 70-74) CIS: controls from the same study area were frequency matched to cases by race and age groups (20-34, 35-44, 45-54, 55-64, and 65-74) NCBCS Phase 3: no controls	Vital status was obtained through linkage to the US National Death Index	N/A	1) Newman, B., Moorman, P. G., Millikan, R., Qaqish, B. F., Geradts, J., Aldrich, T. E., and Liu, E. T. The Carolina Breast Cancer Study: integrating population-based epidemiology and molecular biology. Breast Cancer Res Treat, 35: 51-60, 1995. 2) Millikan, R., Eaton, A., Worley, K., Biscocho, L., Hodgson, E., Huang, W. Y., Geradts, J., Iacocca, M., Cowan, D., Conway, K., and Dressler, L. HER2 codon 655 polymorphism and risk of breast cancer in African Americans and whites. Breast Cancer Res Treat, 79: 355-364, 2003
Oulu Breast Cancer Study	OBCS	Finland	Hospital-based case-control study	Consecutive incident cases diagnosed at the Oulu University Hospital between 2000 and 2004	Healthy, consecutive, anonymous, female Finnish Red-Cross blood donors recruited in 2002 from the same geographical region in Northern Finland	Hospital medical records, Cancer registry, population registry	2012 breast cancer spec/2010 overall	Erkko,H. et al. A recurrent mutation in PALB2 in Finnish cancer families. Nature 446, 316-319 (2007)
Ontario Familial Breast Cancer Registry	OFBCR	Canada	Population-based familial case-control study	Cases diagnosed between 1 Jan 1996-31 Dec 1998 were identified from the Ontario Cancer Registry which registers >97% of all cases residing in the province at the time of diagnosis. All women with invasive breast cancer aged 20–54 years who met the OFBCR definition for high genetic risk (family history of specific cancers particularly breast and ovarian, early onset disease, Ashkenazi ethnicity or a diagnosis of multiple breast cancer) were asked to participate by completing risk factor questionnaires and providing a blood sample. A 25% random sample of individuals in this age category who did not meet the OFBCR definition, 35% of those aged 55–69 at high risk and 8.75% aged 55–69 at low risk were also asked to participate. Individuals diagnosed in 2001 and 2002 were also included if they met high -risk criteria	Unrelated, unaffected population controls were recruited between 2003-2005 by calling randomly selected residential telephone numbers throughout the same geographical region. Eligible controls were women with no history of breast cancer and were frequency-matched by 5-year age group to the expected age distribution of cases. Approximately, 65% of identified eligible women returned questionnaires, and 63% of these donated a blood specimen	Follow-up data including vital status were collected though annual family history follow-up questionnires and though personal history follow-up questionnaires collected at years 10, 15 and 20 since baseline. Vital status of cases was ascertained also through linkage to cancer registry	15-year personal history follow-up information were collected from April 2012 to April 2015	John,E.M. et al. The Breast Cancer Family Registry: an infrastructure for cooperative multinational, interdisciplinary and translational studies of the genetic epidemiology of breast cancer. Breast Cancer Res 6, R375-R389 (2004)
Leiden University Medical Centre Breast Cancer Study	ORIGO	Netherlands	Hospital-based prospective cohort study	Consecutive cases diagnosed 1996-2006 in 2 hospitals of South-West Netherlands (Leiden & Rotterdam). No selection for family history; Rotterdam cases selected for diagnosis aged <70. Cases with in situ carcinomas eligible	Three groups of controls: (1) Blood bank healthy donors from Southwest Netherlands recruited in 1996, 2000 or 2007; (2) People who married a person who was part of a family with high breast cancer risk (BRCA1/2/x). From the Southwest of the Netherlands, recruited 1990-1996; (3) Females tested at the local clinical genetics department for familial diseases, excluding familial cancer syndromes (no mutation found in gene(s) related to the disease being tested), recruited 1995-2007	Linkage to Municipal Population Register; National Pathology Registry; Hospital Information System; General Practitioner	2015 for cases from LUMC (N=890); 2007 for cases from DDHK (N=860)	1) de Bock,G.H. et al. Tumour characteristics and prognosis of breast cancer patients carrying the germline CHEK2*1100delC variant. Journal of Medical Genetics 41, 731-735 (2004). 2) Huijts PE et al . Clinical correlates of low-risk variants in FGFR2, TNRC9, MAP3K1, LSP1 and 8q24 in a Dutch cohort of incident breast cancer cases. Breast Cancer Res.9, R78 (2007)
NCI Polish Breast Cancer Study	PBCS	Poland	Population-based case-control study	Incident cases from 2000-2003 identified through a rapid identification system in participating hospitals covering ~ 90% of all eligible cases, and cancer registries in Warsaw and Lódz covering 100% of all eligible cases	Randomly selected from population lists of all residents of Poland, stratified and frequency matched to cases by case city and age in 5 year categories. Recruited 2000-2003	Complete follow-up for Warsaw cases; Only vital sttus for Lódz cases; We are reviewing medical records plus Cancer Registry data base plus Death Certificates data base	Every 5 years; last data from 2012; currently running follow up after 15 yrs	Garcia-Closas, M. et al. Polymorphisms in DNA double-strand break repair genes and risk of breast cancer: two population-based studies in USA and Poland, and meta-analyses. Hum. Genet. 119, 376-88 (2006)
Karolinska Mammography Project for Risk Prediction of Breast Cancer - Case-Control Study	pKARMA	Sweden	Case-control study	Incident cases from Jan 2001 – Dec 2008 from the Stockholm/Gotland area. Identified through the Stockholm breast cancer registry	Unmatched participants of the KARMA mammography screening study recruited between 2010 and 2011 from Helsingborg and Stockholm	Through the Swedish Cause of Death register, Clinical Breast Cancer register and the inpatient register	We match the data sets to all registers twice a year	Unpublished
Prospective Study of Outcomes in Sporadic Versus Hereditary Breast Cancer	POSH	UK	Prospective cohort	Cases aged 40 or younger at breast cancer diagnosis. Recruited from breast cancer centre oncology clinics across 126 UK hospitals and diagnosed between January 2000 to December 2007	No in-house controls	National data	Vital status updated every 6 months	1) Eccles et al (2007) Prospective study of Outcomes in Sporadic versus Hereditary breast cancer (POSH): Study Protocol. BMC Cancer 7, 160. 2) Tapper W, et al (2008). Association between common genetic variants and prognosis of early onset breast cancer in 30 candidate genes. Breast Cancer Research; 10:R108. 3) Copson E et al (2013) Prospective observational study of breast cancer treatment outcomes for UK women aged 18-40 years at diagnosis: The POSH study. J Natl Cancer Inst 105, 978-988. 4) Copson, E.., et al. (2014). Ethnicity and outcome of young breast cancer patients in the United Kingdom: The POSH study. BR J CANCER 110, 230-241. 5) Copson, E.R., et al. (2015). Obesity and the outcome of young breast cancer patients in the UK: the POSH study. Ann Oncol 26, 101-112. 6) Couch F et al (2015). Inherited mutations in 17 breast cancer susceptibility genes among a large triple-negative breast cancer cohort unselected for family history of breast cancer. J Clin Oncol 33, 304-311
Evaluation of Predictive Factors regarding the Effectivity of Aromatase Inhibitor Therapy	PREFACE	Germany	Multicenter, prospective, randomized, open-label phase IV study	Postmenopausal, steroid hormone receptor positive breast cancer patients who are treated with letrozole. Recruitment at multicentres in Germany between 2009-2011	No controls	Medical records	12/1/2014	No References
Predicting the Risk Of Cancer At Screening Study	PROCAS	UK	Population based study	Women diagnosed with breast cancer since joining the study of women attending the Breast Screening Programme (NHSBSP) in Greater Manchester. Recruitment period Oct 2009-May 2014	Women attending routine NHS breast screening in Greater Manchester without a breast cancer diagnosis. Recruited during the same period as for the cases	Cancer report updated monthly from NHS systems. Deaths report updated weekly from local GP records	Expected to be in follow up until 2020	Evans DG, Astley S, Stavrinos P, Harkness E, Donnelly LS, Dawe S, et al.Improvement in risk prediction, early detection and prevention of breast cancer in the NHS Breast Screening Programme and family history clinics: a dual cohort study. Programme Grants Appl Res 2016;4(11)
Rotterdam Breast Cancer Study	RBCS	Netherlands	Hospital-based case-control study, Rotterdam area	Familial breast cancer patients selected from the Clinical Genetics Center at Erasmus MC Cancer Institute; recruited 1994 - 2005 (RBCS1) and 1995 - 2009 (RBCS2; for OncoArray)	Spouses or mutation-negative siblings of heterozygous Cystic Fibrosis mutation carriers selected from the Clinical Genetics Center at Erasmus MC Cancer Institute; recruited 1996 - 2006 (RBCS1) and 2005 - 2009 (RBCS2)	From medical file/ info from GP or other hospital. Vital status also from Municipal registry	October 2013	M Kriege, A Hollestelle, A Jager et al. Survival and contralateral breast cancer in CHEK2*1100delC breast cancer patients: impact of adjuvant chemotherapy. Br J Cancer 2014; 111(5):1004-13
Singapore and Sweden Breast Cancer Study	SASBAC	Sweden	Population-based case-control study	Incident cases from October 1993 to March 1995 identified via the 6 regional cancer registries in Sweden, to which reporting is mandatory	Controls were randomly selected from the total population registry in 5-year age groups to match the expected age-frequency distribution among cases. Patients and controls were recruited from Oct 1993 through April 1995	Through medical records the first 6-8 years and the Swedish Cause of Death register, Clinical Breast Cancer register and the inpatient register there after	We match the data sets to registers when needed	Wedren, S. et al. Oestrogen receptor alpha gene haplotype and postmenopausal breast cancer risk: a case control study. Breast Cancer Res., 6, R437-49 (2004)
Sheffield Breast Cancer Study	SBCS	UK	Hospital-based case-control study	Women with pathologically confirmed breast cancer recruited from surgical outpatient clinics at the Royal Hallamshire Hospital, Sheffield, 1998 – 2005; cases are a mixture of prevalent and incident disease	Unselected women attending the Sheffield Mammography Screening Service between Sep 2000 - Aug 2004, if their mammograms showed no evidence of a breast lesion	Trent Cancer Registry	9th Sept 2013 (death data)	1) MacPherson,G. et al. Association of a common variant of the CASP8 gene with reduced risk of breast cancer. Journal of the National Cancer Institute 96, 1866-1869 (2004). 2) Rafii,S. et al. A potential role for the XRCC2 R188H polymorphic site in DNA-damage repair and breast cancer. Human Molecular Genetics 11, 1433-1438 (2002)
Study of Epidemiology and Risk factors in Cancer Heredity	SEARCH	UK	Population-based case-control study	2 groups of cases identified through East Anglian Cancer Registry; 1) prevalent cases diagnosed 1991-1996 under 55 years of age at diagnosis, recruited 1996-2002; 2) incident cases diagnosed since 1996 under 70 years of age at diagnosis, recruited 1996-present	Two groups of controls: (1) selected from the EPIC-Norfolk cohort study of 25,000 individuals age 45-74 recruited between 1992 and1994, based in the same geographic region as cases; (2) selected from GP practices from March 2003 to present, frequency matched to cases by age and geographic region	The National Cancer Registration and Analysis Service	8/1/2014	Lesueur, F. et al. Allelic association of the human homologue of the mouse modifier Ptprj with breast cancer. Hum. Mol. Genet., 14, 2349-56 (2005)
Seoul Breast Cancer Study	SEBCS	Korea	Hospital-based case-control study	Consecutive, incident, cases from 2 hospitals in Seoul recruited 2001-2005	Healthy community controls from same catchment area and participating in annual health check-up, 2001-2005	Follow-up information was obtained irregulary as occasion demands	January 2013	1) Lee, K.M. et al. Genetic polymorphisms of ataxia telangiectasia mutated and breast cancer risk. Cancer Epidemiol. Biomarkers Prev., 14, 821-5 (2005). 2) Han,S. et al. CASP8 polymorphisms, estrogen and progesterone receptor status, and breast cancer risk. Breast Cancer Res Treat. 110, 387-393 (2008)
Singapore Breast Cancer Cohort	SGBCC	Singapore	Hospital-based breast cancer cohort and population-based controls	Living breast cancer patients diagnosed with primary insitu or invasive breast cancer at National University Hospital between 2006-2013. Cases are a mixture of prevalent and incident cases	All community-dwelling individuals who are Singaporeans or Singaporean Permanent Residents, 21 years and older. Participants were recruited betwen 2006 and 2010 through word-of-mouth and personal recommendations. In some cases, recruiters also sought participants through "cold-calling" or through door-to-door invitations. Exclusion criteria were a medical history of cancer, acute myocardial infarction or stroke, or major psychiatric morbidity including schizophrenia, psychotic depression, and advanced Alzheimer's Disease	1 visit only, no follow-up required	1 visit only, no follow-up required	No References
Städtisches Klinikum Karlsruhe Deutsches Krebsforschungszentrum Study	SKKDKFZS*	Germany	Hospital-based breast cancer cohort	Women diagnosed with primary in situ or invasive breast cancer at the Städtisches Klinikum Karlsruhe from March 1993 to July 2005	No controls	Follow-up (including vital status) information was obtained from medical records, pathology reports or population registries. Risk factor data were collected from about 10% of patients by questionnaire; from the remaining patients risk factor data were obtained from the medical records	2012 - 2014	Stevens, K.N. et al. 9p13.1 is a triple-negative-specific breast cancer susceptibility locus. Cancer Res. 2012;72(7):1795-803
Swedish Mammography Cohort	SMC	Sweden	Nested case control study from a population-based cohort	All breast cancer cases in the cohort (information from the Swedish Cancer Register) from 1987-2011 for women who gave saliva in 2005-2008 or a blood specimen in 2003-2009 are included	Controls where randomly selected from cancer free women in the cohort who also gave saliva in 2005-2008 or a blood specimen in 2003-2009 . Controls were matched to cases on birth year	Swedish Death Register & Swedish Cancer Register	End of 2014 for mortality, end of 2013 for cancer incidence, end of 2011 for ER/PR receptors	Suzuki R1, Ye W, Rylander-Rudqvist T, Saji S, Colditz GA, Wolk A. Alcohol and postmenopausal breast cancer risk defined by estrogen and progesterone receptor status: a prospective cohort study. J Natl Cancer Inst. 2005 Nov 2;97(21):1601-8

Simultaneous Study of Gemcitabine-Docetaxel Combination adjuvant treatment	SUCCESSB	Germany	Multicenter, prospective, randomized, open-label phase III study	Patients with primary Her2-positive and high risk breast cancer (pN+ oder >pT1b or >G1 or <36y or HR-). Recruitment at multicentres in Germany during 2008-2011	No controls	Medical records	5/1/2014	1) Andermassen U, Neugebauer J, Janni W, Hepp P, Ortmann U, Sommer H, Rack B, Grp SS. (2011) Simultaneous study of gemcitabine-docetaxel combination adjuvant treatment, as well as biological targeted treatment: the SUCCESS B Trial. Breast 20:S66-S66. 2) Jaeger BAS, et al. (2012) HER2 expression on circulating tumor cells (CTC) in patients with early HER2-positive breast cancer: Results of the German SUCCESS B trial. J Clin Oncol 30 (15). 3) Neugebauer JK, et al. (2013) Persistence of HER2 overexpression on circulating tumor cells in patients after systemic treatment for HER2-positive breast cancer: Follow-up results of the German Success B trial. J Clin Oncol 31 (15)
Simultaneous Study of Docetaxel Based Anthracycline Free Adjuvant Treatment Evaluation	SUCCESSC*	Germany	Multicenter, prospective, randomized, open-label phase III study	Patients with primary Her2-negative and high risk breast cancer (pN+ oder >pT1b or >G1 or <36y or HR-). Recruitment at multicentres in Germany during 2009-2011	No controls	Medical records	01/05/2014	1) Hepp P, et al. (2009) Simultaneous Study of Docetaxel Based Anthracycline Free Adjuvant Treatment Evaluation, as Well as Lifestyle Intervention Strategies Success C-Studie. Anticancer Res 29 (5):1567-1568. 2) Jaeger BA, et al. (2015) Persistence of circulating tumor cells immediately after and two years after systemic adjuvant chemotherapy in patients with early breast cancer - Results of the German SUCCESS trials. Cancer Res 75 (9). 3) Ortmann U, Salmen J, Hepp PGM, Beckmann MW, Fehm TN, Hindenburg H, Lichtenegger W, Rack BK, Schneeweiss A, Janni W. (2011) The SUCCESS-C trial: Interim analysis of toxicity evaluating the role of an anthracycline-free chemotherapy regimen in the adjuvant treatment of HER2/neu-negative breast cancer. J Clin Oncol 29 (15). 4) Rack B, et al. (2010) The German SUCCESS C Study - The First European Lifestyle Study on Breast Cancer. Breast Care 5 (6):395-400. doi:10.1159/000322677
IHCC-Szczecin Breast Cancer Study	SZBCS	Poland	Hospital-based case-control study	Prospectively ascertained cases of invasive breast cancer patients diagnosed at the Regional Oncology Hospital (Szczecin) in the years 2002, 2003, 2006 and 2007 or the University Hospital from 2002 to 2007 in Szczecin, West-Pomerania, Poland. Patients with pure intraductal or intralobular cancer were excluded (DCIS or LCIS) but patients with DCIS with micro-invasion were included	Unaffected, matched to cases for year of birth, sex and region; from families with negative cancer family history; controls were part of a population-based study of the 1.3 million inhabitants of West Pomerania performed in 2003 and 2004 designed to identify familial aggregations of cancer by our centre	Questionnaire sent by post	About 10 years ago; currently we are sending questionnaire only to all mutation carriers (BRCA1/2, CHEK2, etc)	1) Jakubowska A, Cybulski C, Szymanska A, Huzarski T, Byrski T, Gronwald J, Debniak T, Górski B, Kowalska E, Narod SA, Lubinski J. BARD1 and breast cancer in Poland. Breast Cancer Res Treat. 2008 Jan;107(1):119-22. 2) Jakubowska A, Jaworska K, Cybulski C, Janicka A, Szymanska-Pasternak J, Lener M, Narod SA, Lubinski J; IHCC-Breast Cancer Study Group. Do BRCA1 modifiers also affect the risk of breast cancer in non-carriers? Eur J Cancer. 2009 Mar;45(5):837-42.3) Cybulski C, Kluzniak W, Huzarski T, Wokolorczyk D, Kashyap A, Jakubowska A, Szwiec M, Byrski T, Debniak T, Górski B, Sopik V, Akbari MR, Sun P, Gronwald J, Narod SA, Lubinski J; Polish Hereditary Breast Cancer Consortium. Clinical outcomes in women with breast cancer and a PALB2 mutation: a prospective cohort analysis. Lancet Oncol. 2015 Jun;16(8):638-44. 4) Cybulski C, Carrot-Zhang J, Kluzniak W, Rivera B, Kashyap A, Wokolorczyk D, Giroux S, Nadaf J, Hamel N, Zhang S, Huzarski T, Gronwald J, Byrski T, Szwiec M, Jakubowska A, Rudnicka H, Lener M, Masojc B, Tonin PN, Rousseau F, Górski B, Debniak T, Majewski J, Lubinski J, Foulkes WD, Narod SA, Akbari MR. Germline RECQL mutations are associated with breast cancer susceptibility. Nat Genet. 2015 Jun;47(6):643-6
Triple Negative Breast Cancer Consortium Study	TNBCC	Various	See studies marked * and 6 studies below for details of individual studies in TNBCC					
Demokritos	DEMOKRITOS	Greece	Hospital-based case-control study	Triple negative breast cancer cases enrolled from 1997-2010 in hospitals serving geographical areas of Greece, including Athens metropolitan area, Thessaloniki, Ioannina, Patras, and Crete (Chania), in collaboration with the Hellenic Cooperative Oncology Group (HECOG).	Regional controls, identified between 2010-2011 from Athens and Thessaloniki, were population-based unaffected women of the same age range.	Not reported	Not reported	Fostira F, et al. Prevalence of BRCA1 mutations among 403 women with triple-negative breast cancer: implications for genetic screening selection criteria: a Hellenic Cooperative Oncology Group Study. Breast Cancer Res Treat [Epub ahead of print] (2012)
A randomized phase III trial comparing nanoparticle-based paclitaxel with solvent-based paclitaxel as part of neoadjuvant chemotherapy for patients with early breast cancer	GEPARSEPTO	Germany	Multicenter, prospective, randomized, open-label phase III study	Patients with early primary breast cancer who are eligible for neoadjuvant chemotherapy. Recruitment at multicentres in Germany during 2011-2013.	No controls	Not reported	Not reported	**1)** Amant F, et al. (2013) Prognosis of Women With Primary Breast Cancer Diagnosed During Pregnancy: Results From an International Collaborative Study. J Clin Oncol 31 (20):2532-+. doi:10.1200/Jco.2012.45.6335. **2)** Gerber B, et al. (2014) Surgical Outcome after Neoadjuvant Chemotherapy and Bevacizumab: Results from the GeparQuinto Study (GBG 44). Ann Surg Oncol 21 (8):2517-2524. doi:10.1245/s10434-014-3606-9. **3)** Jackisch C,et al. (2013) A randomized phase III trial comparing nanoparticle-based paclitaxel with solvent-based paclitaxel as part of neoadjuvant chemotherapy for patients with early breast cancer (GeparSepto): GBG 69. J Clin Oncol 31 (15). **4)** Untch M, et al. (2015) A randomized phase III trial comparing neoadjuvant chemotherapy with weekly nanoparticle-based paclitaxel with solvent-based paclitaxel followed by anthracycline/cyclophosphamide for patients with early breast cancer (GeparSepto): GBG 69. Cancer Res 75 (9). doi:10.1158/1538-7445.SABCS14-S2-07.
University of Kansas Medical Center	KUMC	USA				Not reported	Not reported	No References
Ohio State University	OSU	USA	Hospital-based case-control study	Incident triple negative invasive breast cancer cases from a collection of incident breast cancer cases diagnosed in Columbus, Ohio (2006-2011).	Population-matched controls accrued through primary care clinics in the OSU medical center system (2006-2011).	Not reported	Not reported	No references yet
Roswell Park Cancer Institute	RPCI	USA	Hospital-based case-control study	Triple negative invasive breast cancer cases from incident cases recruited to the RPCI Data Bank and Biorepository from 2006-2010.	Healthy controls identified from employee volunteers, and women recruited from community events from 2006-2010.	Not reported	Not reported	Ambrosone CB, et al. Establishing a cancer center data bank and biorepository for multidisciplinary research. Cancer Epidemiol Biomarkers Prev. 15(9):1575-7 (2006)
University of Texas MD Anderson Cancer Center	UTMDACC	USA				Not reported	Not reported	No References
Taiwanese Breast Cancer Study	TWBCS	Taiwan	Hospital-based case-control study	Incident cases diagnosed & treated at 2 major teaching hospitals in Taiwan. [between March 2002 and August 2005]	Controls cancer-free individuals, randomly selected from women attending health exam. at same hospital during study period. Underwent 1-day health examination - any showing evidence cancer excluded	The information was from the cooperation of the hospitals	2013	1) Hsu, HM et al. Breast cancer risk is associated with genes encoding the DNA double-strand break repair Mre11/Rad50/Nbs1 complex. Cancer Epidemiol. Biomarkers Prev. 16, 2024-32 (2007). 2) Ding, SI, et al. Genetic variants of BLM interact with RAD51 to increase breast cancer susceptibility. Carcinogenesis. 30, 43-9 (2009)
UCI Breast Cancer Study	UCIBCS*	USA	Population-based case-control study	All cases diagnosed in Orange County, California, during one-year period beginning March 1, 1994. Ascertained through the population-based Cancer Surveillance Program of Orange County California (CSPOC)	Female controls under age 75 years without history of cancer recruited using random digit dialing among Orange County residents & frequency matched to cases by age & race/ethnicity. Recruited from 1998-2003	The vital status and follow-up date are standard items of the California Cancer Registry	Completed information is as of 10/30/2013	1) Anton-Culver, H.et al. Characteristics of BRCA1 mutations in a population-based case series of breast and ovarian cancer. Eur. J. Cancer, 36, 1200-8 (2000). 2) Ziogas, A. et al. Cancer risk estimates for family members of a population-based family registry for breast and ovarian cancer. Cancer Epidemiol. Biomarkers Prev., 9, 103-11 (2000)
	*Subset of samples are part of the TNBCC							

Variant (N=313)[a]	Chromosome	Position	Reference Allele	Effect Allele	Overall Breast Cancer (coefficient)	ER positive (coefficient)[b]	ER negative (coefficient)[b]	Imputation quality r^2 OncoArray (European)[c]	Imputation quality r^2 iCOGS (European)[c]	Imputation quality r^2 OncoArray (Asian)[c]	Imputation quality r^2 iCOGS (Asian)[c]
1_100880328_A_T	1	100880328	A	T	0.0373	0.0373	0.0373	1.0000	0.9200	1.0000	0.9060
1_10566215_A_G	1	10566215	A	G	-0.0586	-0.0407	-0.1109	1.0000	1.0000	1.0000	1.0000
1_110198129_CAAA_C	1	110198129	CAAA	C	0.0458	0.0458	0.0458	0.9400	0.8590	0.6880	0.5610
1_114445880_G_A	1	114445880	G	A	0.0621	0.0621	0.0621	0.9980	0.9980	0.9910	0.9980
1_118141492_A_C	1	118141492	A	C	0.0452	0.0452	0.0452	0.9970	0.9780	0.9920	0.9580
1_120257110_T_C	1	120257110	T	C	0.0385	0.0430	0.0226	0.9960	1.0000	0.9980	1.0000
1_121280613_A_G	1	121280613	A	G	0.0881	0.1052	0.0209	1.0000	1.0000	1.0000	1.0000
1_121287994_A_G	1	121287994	A	G	-0.0673	-0.0814	-0.0114	0.9620	0.9560	0.9770	0.9880
1_145604302_C_CT	1	145604302	C	CT	-0.0399	-0.0469	-0.0126	0.9380	0.9080	0.9360	0.8800
1_149906413_T_C	1	149906413	T	C	0.0548	0.0548	0.0548	1.0000	0.9720	1.0000	0.9790
1_155556971_G_A	1	155556971	G	A	0.0499	0.0499	0.0499	0.9950	0.9800	0.9430	0.9330
1_168171052_CA_C	1	168171052	CA	C	-0.0680	-0.0680	-0.0680	0.9000	0.7070	0.5120	0.4880
1_172328767_T_TA	1	172328767	T	TA	-0.0435	-0.0435	-0.0435	0.9200	0.8860	0.8450	0.8150
1_18807339_T_C	1	18807339	T	C	-0.0564	-0.0649	-0.0248	0.9990	0.8670	0.9960	0.8380
1_201437832_C_T	1	201437832	C	T	0.0917	0.0917	0.0917	1.0000	0.8200	1.0000	0.4860
1_202184600_C_T	1	202184600	C	T	-0.0065	0.0133	-0.0822	1.0000	0.9810	1.0000	0.9670
1_203770448_T_A	1	203770448	T	A	0.0498	0.0498	0.0498	0.9980	0.9930	0.9990	0.9920
1_204502514_T_TTCTGAAA	1	204502514	T	TTCTGAAACAGGC	-0.0321	-0.0024	-0.1345	0.9720	0.9430	0.6770	0.6350
1_208076291_G_A	1	208076291	G	A	-0.0366	-0.0366	-0.0366	0.9950	0.9640	0.9710	0.9390
1_217053815_T_G	1	217053815	T	G	0.0417	0.0417	0.0417	0.9030	0.7020	0.9180	0.5890
1_217220574_G_A	1	217220574	G	A	-0.0440	-0.0459	0.0029	0.9960	0.9990	0.9030	0.9570
1_220671050_C_T	1	220671050	C	T	0.0418	0.0418	0.0418	0.9520	0.8820	0.9550	0.8980
1_242034263_A_G	1	242034263	A	G	0.1428	0.1428	0.1428	1.0000	0.9900	1.0000	0.8860
1_41380440_C_T	1	41380440	C	T	0.0426	0.0426	0.0426	0.9840	0.7510	0.9760	0.5460
1_41389220_T_C	1	41389220	T	C	0.1550	0.1550	0.1550	0.9620	0.8450	0.9170	0.7310
1_46670206_TC_T	1	46670206	TC	T	0.0447	0.0595	0.0216	1.0000	0.9540	1.0000	0.9120
1_51467096_CT_C	1	51467096	CT	C	0.0374	0.0374	0.0374	0.9020	0.8890	0.8070	0.8080
1_7917076_G_A	1	7917076	G	A	-0.0409	-0.0409	-0.0409	0.9910	0.7920	0.9710	0.6820
1_88156923_G_A	1	88156923	G	A	0.0494	0.0580	0.0183	1.0000	1.0000	1.0000	1.0000
1_88428199_C_A	1	88428199	C	A	-0.0387	-0.0387	-0.0387	1.0000	1.0000	1.0000	1.0000
2_10138983_T_C	2	10138983	T	C	0.0603	0.0603	0.0603	0.9580	0.9340	0.7470	0.7480
2_121058254_A_G	2	121058254	A	G	-0.0334	-0.0232	-0.0682	0.9960	0.9990	0.9980	1.0000
2_121089731_T_C	2	121089731	T	C	-0.0427	-0.0290	-0.1027	0.9260	1.0000	0.8480	1.0000
2_121159205_G_A	2	121159205	G	A	-0.0440	-0.0440	-0.0440	1.0000	0.8630	1.0000	0.7270
2_121246568_T_C	2	121246568	T	C	0.0992	0.0992	0.0992	1.0000	0.9970	1.0000	0.9930
2_172974566_C_G	2	172974566	C	G	-0.0473	-0.0611	-0.0061	1.0000	1.0000	1.0000	1.0000
2_174212910_A_G	2	174212910	A	G	0.0593	0.0621	0.0175	1.0000	0.9960	1.0000	0.9790
2_192381934_C_T	2	192381934	C	T	0.0316	0.0180	0.1012	1.0000	0.5300	1.0000	0.4950
2_19315675_T_A	2	19315675	T	A	-0.0331	-0.0229	-0.0570	1.0000	0.9960	1.0000	0.9910
2_202204741_T_C	2	202204741	T	C	-0.0492	-0.0492	-0.0492	1.0000	1.0000	1.0000	1.0000
2_217920769_G_T	2	217920769	G	T	-0.1318	-0.1532	-0.0589	1.0000	1.0000	1.0000	1.0000
2_217955896_GA_G	2	217955896	GA	G	-0.2016	-0.2362	-0.0558	0.9840	0.9040	0.9580	0.7890
2_218292158_C_G	2	218292158	C	G	-0.0757	-0.0757	-0.0757	0.9770	0.9730	0.9870	0.9900
2_218714845_G_A	2	218714845	G	A	-0.0431	-0.0463	-0.0184	1.0000	0.6770	1.0000	0.7350
2_241388857_C_A	2	241388857	C	A	-0.1232	-0.1232	-0.1232	1.0000	0.7360	1.0000	0.4340
2_25129473_A_G	2	25129473	A	G	-0.0427	-0.0427	-0.0427	1.0000	0.7080	0.9990	0.7750
2_29179452_G_C	2	29179452	G	C	-0.0066	0.0207	-0.1006	1.0000	0.9890	1.0000	0.9730
2_29615233_T_C	2	29615233	T	C	-0.0427	-0.0427	-0.0427	0.9840	0.7710	0.9100	0.5720
2_39699510_C_CT	2	39699510	C	CT	-0.0402	-0.0402	-0.0402	0.9020	0.7580	0.6360	0.5350
2_70172587_G_A	2	70172587	G	A	-0.0412	-0.0412	-0.0412	0.9660	0.8050	0.9490	0.7930
2_88358825_G_C	2	88358825	G	C	0.0473	0.0473	0.0473	0.9430	0.5970	0.9490	0.5520
3_141112859_CTT_C	3	141112859	CTT	C	0.0551	0.0551	0.0551	0.9970	0.9940	0.9900	0.9620
3_172285237_G_A	3	172285237	G	A	0.0422	0.0501	-0.0133	0.9970	0.9970	0.9820	0.9830
3_189774456_C_T	3	189774456	C	T	-0.0478	-0.0478	-0.0478	0.9960	0.7060	0.9610	0.5980
3_27353716_C_A	3	27353716	C	A	0.0748	0.0822	0.0310	1.0000	1.0000	1.0000	1.0000
3_27388664_C_G	3	27388664	C	G	0.0502	0.0502	0.0502	1.0000	1.0000	1.0000	1.0000
3_29294845_C_T	3	29294845	C	T	-0.1281	-0.1221	-0.2988	0.9100	0.8480	0.2350	0.3490
3_30684907_C_T	3	30684907	C	T	0.0592	0.0657	0.0170	1.0000	1.0000	1.0000	1.0000
3_46888198_T_C	3	46888198	T	C	-0.0806	-0.0806	-0.0806	0.9930	0.9280	0.9940	0.8990
3_4742251_A_G	3	4742251	A	G	0.0616	0.0616	0.0616	1.0000	0.9990	1.0000	0.9970
3_49709912_C_CT	3	49709912	C	CT	-0.0367	-0.0355	-0.0721	0.9470	0.9370	0.8670	0.8190
3_55970777_A_AT	3	55970777	A	AT	-0.1195	-0.1195	-0.1195	0.9600	0.5630	0.7540	0.4500
3_59373745_C_T	3	59373745	C	T	-0.0394	-0.0394	-0.0394	0.9970	0.6400	0.9740	0.5500
3_63887449_T_TTG	3	63887449	T	TTG	0.0648	0.0648	0.0648	0.9880	0.9820	0.9980	0.9960
3_71620370_T_G	3	71620370	T	G	-0.0374	-0.0374	-0.0374	0.9860	0.8860	0.9880	0.8570
3_87037543_A_G	3	87037543	A	G	-0.0723	-0.0723	-0.0723	0.9430	1.0000	0.6670	1.0000
3_99403877_G_A	3	99403877	G	A	-0.0376	-0.0376	-0.0376	0.9950	0.9780	0.9930	0.9750
4_106069013_G_T	4	106069013	G	T	0.0471	0.0594	0.0097	1.0000	0.9980	1.0000	0.9930
4_126752992_A_AAT	4	126752992	A	AAT	-0.0377	-0.0377	-0.0377	0.9630	0.9640	0.9500	0.9310
4_143467195_C_T	4	143467195	C	T	-0.0569	-0.0569	-0.0569	0.9990	0.8140	0.9970	0.5800
4_151218296_CATATTT_C	4	151218296	CATATTT	C	0.0388	0.0388	0.0388	0.9910	0.9830	0.9830	0.9820
4_175842495_G_A	4	175842495	G	A	-0.0898	-0.1162	0.0199	1.0000	1.0000	1.0000	0.9940
4_175847436_C_A	4	175847436	C	A	0.0348	0.0537	-0.0099	1.0000	0.9830	1.0000	0.9700
4_187503758_A_T	4	187503758	A	T	0.0357	0.0357	0.0357	0.9970	0.9700	0.9940	0.9730
4_38784633_G_T	4	38784633	G	T	0.0489	0.0489	0.0489	0.9990	1.0000	0.9980	1.0000
4_84370124_TAA_TA	4	84370124	TAA	TA	-0.0464	-0.0464	-0.0464	0.9440	0.9390	0.9560	0.9560
4_89240476_G_A	4	89240476	G	A	0.0352	0.0352	0.0352	0.9740	1.0000	0.9420	1.0000
4_92594859_TTCTTTC_T	4	92594859	TTCTTTC	T	-0.0407	-0.0407	-0.0407	0.9400	0.7990	0.9320	0.8320
5_104300273_G_T	5	104300273	G	T	-0.0487	-0.0487	-0.0487	0.9940	0.8240	0.9860	0.7320
5_122478676_C_A	5	122478676	C	A	-0.0386	-0.0386	-0.0386	0.9990	0.9790	0.9990	0.9910
5_122705244_C_T	5	122705244	C	T	0.0944	0.0944	0.0944	0.9970	1.0000	0.9870	0.9940
5_1279790_C_T	5	1279790	C	T	0.0617	0.0325	0.1502	1.0000	1.0000	1.0000	1.0000
5_1296255_A_AG	5	1296255	A	AG	-0.0549	-0.0417	-0.1056	1.0000	0.9960	1.0000	0.9880
5_131640536_A_G	5	131640536	A	G	0.0392	0.0467	0.0099	0.9810	0.9830	0.9550	0.9550
5_132407058_C_T	5	132407058	C	T	-0.0388	-0.0561	-0.0214	0.9990	0.7640	0.9980	0.7980
5_1353077_T_C	5	1353077	T	C	0.1552	0.1552	0.1552	1.0000	0.9030	1.0000	0.2850
5_158244083_C_T	5	158244083	C	T	-0.0677	-0.0677	-0.0677	1.0000	1.0000	1.0000	1.0000
5_16231194_G_C	5	16231194	G	C	-0.0426	-0.0426	-0.0426	1.0000	0.9990	1.0000	0.9980
5_169591460_T_C	5	169591460	T	C	0.0412	0.0501	0.0182	0.9970	0.9150	0.9980	0.9550
5_173358154_G_A	5	173358154	G	A	0.0365	0.0365	0.0365	0.9870	0.9920	0.9900	0.9920
5_176134882_T_C	5	176134882	T	C	0.0363	0.0363	0.0363	0.9950	0.7080	0.9970	0.6640
5_2777029_G_A	5	2777029	G	A	0.0391	0.0391	0.0391	0.9840	0.9850	0.9670	0.9700
5_32579616_TCA_T	5	32579616	TCA	T	0.0363	0.0363	0.0363	1.0000	0.9190	0.9980	0.9590
5_345109_T_C	5	345109	T	C	0.0840	0.0840	0.0840	0.9850	0.9710	0.9730	0.9010
5_44508264_G_GT	5	44508264	G	GT	-0.1177	-0.1177	-0.1177	0.9700	0.9750	0.6000	0.7460
5_44619502_A_G	5	44619502	A	G	-0.1101	-0.1101	-0.1101	0.9910	0.9850	0.9850	0.9860
5_44649944_C_T	5	44649944	C	T	0.0492	0.0713	-0.0261	1.0000	1.0000	1.0000	1.0000
5_44706498_A_G	5	44706498	A	G	0.0497	0.0648	-0.0256	1.0000	1.0000	1.0000	1.0000
5_44853593_G_C	5	44853593	G	C	-0.0336	-0.0222	-0.0778	1.0000	1.0000	1.0000	1.0000
5_52679539_C_CA	5	52679539	C	CA	0.0571	0.0571	0.0571	0.9720	0.9310	0.9660	0.9480
5_55662540_C_CT	5	55662540	C	CT	-0.0458	-0.0458	-0.0458	0.9640	0.7610	0.9650	0.7320
5_55965167_C_T	5	55965167	C	T	0.0394	0.0394	0.0394	1.0000	0.7280	1.0000	0.6840
5_56023083_T_G	5	56023083	T	G	0.1366	0.1612	0.0686	1.0000	1.0000	1.0000	1.0000
5_56042972_C_T	5	56042972	C	T	0.0865	0.1082	0.0058	0.9990	0.9990	0.9990	0.9990
5_56045081_T_C	5	56045081	T	C	-0.0564	-0.0643	-0.0168	1.0000	1.0000	1.0000	1.0000
5_58241712_C_T	5	58241712	C	T	-0.0434	-0.0434	-0.0434	0.9470	0.8200	0.9420	0.8190
5_71965007_G_A	5	71965007	G	A	-0.0410	-0.0410	-0.0410	0.9450	0.8720	0.9590	0.8520
5_73234583_T_C	5	73234583	T	C	-0.0363	-0.0494	-0.0101	0.9810	0.8180	0.9580	0.8340
5_77155397_GT_G	5	77155397	GT	G	-0.0408	-0.0408	-0.0408	0.9940	0.6600	0.9960	0.7020
5_79180995_G_GA	5	79180995	G	GA	0.0328	0.0248	0.0804	0.9880	0.9910	0.9920	0.9830
5_81512947_TA_T	5	81512947	TA	T	-0.0598	-0.0731	-0.0342	1.0000	0.9630	1.0000	0.8860
5_90789470_G_A	5	90789470	G	A	-0.0564	-0.0714	-0.0031	1.0000	0.7910	1.0000	0.8650
6_130341728_C_CT	6	130341728	C	CT	0.0472	0.0472	0.0472	0.9820	0.9720	0.9710	0.9610
6_13713366_G_C	6	13713366	G	C	-0.0553	-0.0623	-0.0152	1.0000	0.9750	1.0000	0.9670
6_149595505_T_C	6	149595505	T	C	-0.0476	-0.0476	-0.0476	0.9960	0.9590	0.9920	0.9700
6_151949806_A_C	6	151949806	A	C	0.0703	0.0541	0.1103	1.0000	1.0000	1.0000	1.0000
6_151955914_A_G	6	151955914	A	G	0.1449	0.1150	0.2240	1.0000	1.0000	1.0000	1.0000
6_152022664_CAAAAAAA_C	6	152022664	CAAAAAAA	C	0.0137	0.0137	0.0137	0.9200	0.8900	0.8160	0.8210
6_152023191_G_A	6	152023191	G	A	0.0626	0.0509	0.1008	1.0000	1.0000	1.0000	1.0000
6_152055978_A_T	6	152055978	A	T	0.0740	0.0740	0.0740	0.9940	0.9950	0.9770	0.9800
6_152432902_C_T	6	152432902	C	T	0.0649	0.0527	0.0965	1.0000	0.9970	1.0000	0.9970
6_16399557_C_T	6	16399557	C	T	-0.0373	-0.0373	-0.0373	1.0000	0.7070	1.0000	0.6770
6_169006947_C_G	6	169006947	C	G	-0.0308	-0.0252	-0.0628	0.9470	0.9320	0.8280	0.7960
6_170332621_T_C	6	170332621	T	C	0.0373	0.0373	0.0373	0.9920	0.9690	0.9600	0.9500
6_18783140_G_A	6	18783140	G	A	0.0326	0.0478	0.0033	0.9980	0.8780	0.9980	0.8190
6_20537845_CA_C	6	20537845	CA	C	-0.0391	-0.0391	-0.0391	0.9080	0.8860	0.8970	0.8760
6_21923810_T_C	6	21923810	T	C	-0.0321	-0.0438	-0.0032	0.9990	0.9970	0.9980	0.9970
6_27425644_G_C	6	27425644	G	C	-0.0737	-0.0737	-0.0737	0.9910	0.9420	0.0000	0.6150
6_43227141_G_A	6	43227141	G	A	-0.0640	-0.0640	-0.0640	0.9950	0.9300	0.9700	0.8100
6_82263549_AAT_A	6	82263549	AAT	A	0.0477	0.0477	0.0477	0.9340	0.9080	0.9370	0.9220
6_85912194_CAA_C	6	85912194	CAA	C	0.0762	0.0762	0.0762	0.9610	0.9020	0.6370	0.4350
6_87803819_T_C	6	87803819	T	C	0.0383	0.0318	0.0678	0.9890	0.8830	0.9710	0.8580
7_101552440_G_A	7	101552440	G	A	-0.0568	-0.0568	-0.0568	0.9940	0.8220	0.9920	0.7460
7_102481842_T_C	7	102481842	T	C	0.0418	0.0418	0.0418	0.9970	0.9910	0.9930	0.9910
7_130656911_C_T	7	130656911	C	T	-0.0476	-0.0476	-0.0476	0.9520	0.9290	0.9020	0.8610
7_130674481_G_A	7	130674481	G	A	0.0416	0.0416	0.0416	1.0000	0.8740	1.0000	0.8640
7_139943702_CT_C	7	139943702	CT	C	0.0582	0.0666	0.0057	0.9340	0.8580	0.9680	0.8590
7_144048902_G_T	7	144048902	G	T	-0.0563	-0.0592	-0.0148	0.9120	0.8960	0.9070	0.9110
7_21940960_A_G	7	21940960	A	G	-0.0467	-0.0467	-0.0467	0.9900	1.0000	0.9910	1.0000
7_25569548_C_T	7	25569548	C	T	-0.0486	-0.0486	-0.0486	0.9650	0.8790	0.8360	0.7330
7_28869017_G_A	7	28869017	G	A	-0.0572	-0.0572	-0.0572	0.9930	0.9960	0.9420	0.9730
7_55192256_A_C	7	55192256	A	C	-0.0349	-0.0349	-0.0349	1.0000	1.0000	1.0000	1.0000

7_91459189_A_ATT	7	91459189	A	ATT	0.0452	0.0452	0.0452	0.9200	0.9310	0.9560	0.9650
7_94113799_T_C	7	94113799	T	C	0.0449	0.0449	0.0449	0.9950	0.9940	0.9990	0.9950
7_98005235_G_A	7	98005235	G	A	-0.0467	-0.0467	-0.0467	1.0000	1.0000	1.0000	1.0000
7_99948655_T_G	7	99948655	T	G	0.0420	0.0420	0.0420	0.9840	0.8900	0.9720	0.7640
8_102483100_T_C	8	102483100	T	C	0.0593	0.0736	0.0137	0.9500	0.9440	0.6000	0.8270
8_106358620_A_T	8	106358620	A	T	-0.0745	-0.0895	-0.0100	0.9790	0.7560	0.9810	0.6720
8_117209548_A_G	8	117209548	A	G	-0.0417	-0.0417	-0.0417	1.0000	0.9820	1.0000	0.9910
8_120862186_A_G	8	120862186	A	G	0.0527	0.0527	0.0527	0.9750	0.9680	0.9770	0.9670
8_124563705_T_C	8	124563705	T	C	0.0477	0.0477	0.0477	0.9900	0.9210	0.9840	0.8650
8_124571581_G_A	8	124571581	G	A	0.0340	0.0340	0.0340	0.9910	0.9550	0.9870	0.9680
8_124739913_T_G	8	124739913	T	G	0.0466	0.0466	0.0466	0.9810	0.9840	0.9570	0.9740
8_128213561_C_CA	8	128213561	C	CA	-0.0430	-0.0430	-0.0430	1.0000	0.8550	1.0000	0.7370
8_128370949_C_G	8	128370949	C	G	0.0642	0.0820	0.0076	0.9990	0.9990	0.9990	0.9990
8_128372172_A_G	8	128372172	A	G	0.0597	0.0597	0.0597	1.0000	1.0000	0.9980	0.9980
8_129199566_G_A	8	129199566	G	A	0.0615	0.0615	0.0615	1.0000	0.9970	1.0000	0.9980
8_143669254_A_G	8	143669254	A	G	-0.0346	-0.0346	-0.0346	0.9600	0.7790	0.9040	0.7900
8_170692_T_C	8	170692	T	C	0.0477	0.0348	0.1040	0.9230	0.7450	0.9590	0.7950
8_17787610_CT_C	8	17787610	CT	C	-0.0377	-0.0377	-0.0377	0.9310	0.8840	0.8370	0.7520
8_23447496_A_G	8	23447496	A	G	-0.0389	-0.0389	-0.0389	1.0000	1.0000	1.0000	1.0000
8_23663653_C_A	8	23663653	C	A	0.0335	0.0451	0.0059	0.9990	0.9990	0.9990	0.9980
8_29509616_A_C	8	29509616	A	C	-0.0601	-0.0601	-0.0601	1.0000	1.0000	1.0000	1.0000
8_36858483_A_G	8	36858483	A	G	-0.0760	-0.0760	-0.0760	1.0000	0.9910	1.0000	0.9890
8_76230943_A_G	8	76230943	A	G	0.0755	0.0755	0.0755	1.0000	1.0000	0.9970	0.9950
8_76333056_C_T	8	76333056	C	T	0.1129	0.1129	0.1129	1.0000	0.9970	1.0000	0.9800
8_76378165_G_T	8	76378165	G	T	-0.0391	-0.0391	-0.0391	1.0000	0.9400	1.0000	0.9170
9_110303808_TAA_T	9	110303808	TAA	T	0.0797	0.1007	0.0130	0.9920	0.9120	0.9950	0.9400
9_110837073_A_G	9	110837073	A	G	0.1158	0.1315	0.0289	1.0000	1.0000	1.0000	1.0000
9_110837176_C_T	9	110837176	C	T	0.0653	0.0809	-0.0037	1.0000	1.0000	1.0000	1.0000
9_110849525_G_T	9	110849525	G	T	0.0153	0.0153	0.0153	1.0000	0.8920	1.0000	0.6520
9_110885479_C_T	9	110885479	C	T	0.0877	0.1110	0.0019	1.0000	1.0000	1.0000	1.0000
9_119313486_A_G	9	119313486	A	G	-0.0462	-0.0462	-0.0462	0.9880	0.9820	0.9830	0.9700
9_129424719_A_G	9	129424719	A	G	-0.0382	-0.0382	-0.0382	0.9560	0.9030	0.9080	0.8690
9_136146597_C_T	9	136146597	C	T	0.0400	0.0400	0.0400	0.9990	0.9860	0.9990	0.9820
9_21964882_CAAAA_C	9	21964882	CAAAA	C	0.0550	0.0550	0.0550	0.9710	0.9650	0.8960	0.9020
9_22041998_C_G	9	22041998	C	G	0.0289	0.0168	0.0906	1.0000	1.0000	1.0000	1.0000
9_36928288_T_C	9	36928288	T	C	0.0249	0.0249	0.0249	0.9940	0.9800	0.9860	0.9720
9_6880263_A_G	9	6880263	A	G	0.0348	0.0499	-0.0078	1.0000	0.8470	1.0000	0.6880
9_87782211_T_C	9	87782211	T	C	0.0361	0.0361	0.0361	1.0000	1.0000	1.0000	1.0000
9_98362587_T_C	9	98362587	T	C	0.0576	0.0576	0.0576	0.9920	0.9790	0.9810	0.9860
10_114777670_C_T	10	114777670	C	T	0.0472	0.0472	0.0472	1.0000	0.9990	1.0000	0.9990
10_115128491_T_C	10	115128491	T	C	-0.0592	-0.0592	-0.0592	1.0000	1.0000	1.0000	1.0000
10_123095209_G_A	10	123095209	G	A	-0.0538	-0.0702	0.0048	0.9970	0.9640	0.9940	0.9500
10_123340107_A_G	10	123340107	A	G	0.1508	0.1837	0.0053	0.9980	0.9830	0.9980	0.9860
10_123340431_GC_G	10	123340431	GC	G	-0.2408	-0.2913	-0.0326	0.9990	0.9980	0.9990	0.9950
10_123349324_A_T	10	123349324	A	T	-0.2609	-0.3270	-0.0137	0.9660	0.9380	0.5830	0.6880
10_13892298_G_A	10	13892298	G	A	0.0371	0.0371	0.0371	1.0000	0.9740	1.0000	0.9390
10_22032942_A_G	10	22032942	A	G	-0.0580	-0.0719	0.0344	1.0000	1.0000	1.0000	1.0000
10_22477776_ACC_A	10	22477776	ACC	A	0.1687	0.1687	0.1687	0.9930	0.9930	0.9810	0.9650
10_22861490_A_C	10	22861490	A	C	0.0875	0.0960	0.0201	0.9770	0.9740	0.9440	0.9490
10_38523626_C_A	10	38523626	C	A	0.0404	0.0404	0.0404	0.9510	0.9410	0.9940	0.9890
10_5794652_A_G	10	5794652	A	G	0.0470	0.0470	0.0470	1.0000	1.0000	1.0000	1.0000
10_64299890_A_G	10	64299890	A	G	-0.1345	-0.1428	-0.1030	0.9830	0.9820	0.9940	0.9920
10_64819996_G_T	10	64819996	G	T	0.0472	0.0472	0.0472	1.0000	1.0000	1.0000	1.0000
10_71335574_C_T	10	71335574	C	T	-0.0404	-0.0404	-0.0404	0.9580	0.7060	0.9060	0.5620
10_80851257_G_T	10	80851257	G	T	-0.0805	-0.0898	-0.0443	1.0000	1.0000	1.0000	1.0000
10_80886726_A_G	10	80886726	A	G	0.0762	0.0762	0.0762	1.0000	0.9970	1.0000	0.9890
10_95292187_CAA_C	10	95292187	CAA	C	-0.0512	-0.0512	-0.0512	0.9430	0.8730	0.9000	0.8380
11_103614438_T_G	11	103614438	T	G	0.0147	0.0029	0.0676	0.9920	0.7650	0.9960	0.6630
11_108267402_C_CA	11	108267402	C	CA	-0.0022	0.0141	-0.0629	0.9980	0.9940	1.0000	0.9960
11_111696440_T_C	11	111696440	T	C	-0.0396	-0.0396	-0.0396	0.9980	0.9980	0.9980	0.9990
11_116727936_A_T	11	116727936	A	T	-0.0423	-0.0423	-0.0423	0.9980	0.9970	0.9990	0.9980
11_122966626_A_G	11	122966626	A	G	-0.0383	-0.0383	-0.0383	0.9980	0.9080	0.9970	0.8500
11_129243417_T_G	11	129243417	T	G	-0.0543	-0.0543	-0.0543	0.9900	0.9210	0.9890	0.9230
11_129461016_A_G	11	129461016	A	G	0.0453	0.0453	0.0453	1.0000	0.9990	1.0000	0.9930
11_18664241_T_G	11	18664241	T	G	0.0461	0.0461	0.0461	0.9550	0.7170	0.9400	0.7390
11_1895708_C_A	11	1895708	C	A	-0.0762	-0.0762	-0.0762	1.0000	0.9870	1.0000	0.9880
11_42844441_C_T	11	42844441	C	T	-0.0336	-0.0336	-0.0336	1.0000	0.8070	1.0000	0.7670
11_433617_T_C	11	433617	T	C	-0.0437	-0.0437	-0.0437	0.9800	0.8400	0.9340	0.7410
11_44368892_G_A	11	44368892	G	A	0.0374	0.0374	0.0374	0.9930	0.9890	0.9940	0.9930
11_46318032_C_G	11	46318032	C	G	-0.0748	-0.0748	-0.0748	0.9280	0.8000	0.3280	0.4530
11_65553492_C_A	11	65553492	C	A	0.0425	0.0425	0.0425	0.9970	0.9870	0.9980	0.9770
11_65572431_G_A	11	65572431	G	A	-0.0347	-0.0448	-0.0067	1.0000	0.9900	0.9990	0.9960
11_69328130_A_T	11	69328130	A	T	-0.0423	-0.0538	0.0143	0.9540	0.9610	0.9300	0.9360
11_69330983_G_A	11	69330983	G	A	0.1022	0.1240	0.0174	1.0000	0.9990	0.9990	0.9990
11_69331418_C_T	11	69331418	C	T	0.1782	0.2018	0.0066	0.9910	0.9910	0.8950	0.9390
11_803017_A_G	11	803017	A	G	0.0457	0.0457	0.0457	0.9940	0.9930	0.9970	1.0000
12_103097887_C_T	12	103097887	C	T	0.0546	0.0611	0.0149	0.9910	0.9570	0.8740	0.8190
12_111600134_G_T	12	111600134	G	T	-0.0442	-0.0442	-0.0442	1.0000	0.8830	1.0000	0.7300
12_115108136_T_C	12	115108136	T	C	0.0465	0.0465	0.0465	1.0000	0.9700	1.0000	0.9310
12_115796577_A_G	12	115796577	A	G	-0.0428	-0.0643	-0.0148	0.9970	0.9990	0.9900	0.9950
12_115835836_T_C	12	115835836	T	C	-0.0813	-0.0977	-0.0153	1.0000	1.0000	1.0000	1.0000
12_120832146_C_T	12	120832146	C	T	0.0516	0.0516	0.0516	1.0000	0.8090	1.0000	0.6140
12_14413931_G_C	12	14413931	G	C	0.0484	0.0484	0.0484	1.0000	1.0000	1.0000	1.0000
12_28149568_C_T	12	28149568	C	T	-0.0620	-0.0620	-0.0620	1.0000	1.0000	1.0000	1.0000
12_28174817_C_T	12	28174817	C	T	-0.0856	-0.0856	-0.0856	1.0000	1.0000	1.0000	1.0000
12_28347382_C_T	12	28347382	C	T	-0.0521	-0.0521	-0.0521	0.9760	0.9810	0.8190	0.9360
12_29140260_G_A	12	29140260	G	A	0.0647	0.0647	0.0647	0.9980	0.9410	0.9920	0.9170
12_293626_A_G	12	293626	A	G	0.0401	0.0401	0.0401	0.9950	0.6100	0.9950	0.6350
12_57146069_T_G	12	57146069	T	G	-0.0579	-0.0579	-0.0579	1.0000	1.0000	1.0000	1.0000
12_70798355_A_T	12	70798355	A	T	0.0469	0.0469	0.0469	0.9960	0.9330	0.9870	0.9910
12_83064195_G_GA	12	83064195	G	GA	0.0671	0.0671	0.0671	0.9700	0.4990	0.9390	0.4820
12_85004551_C_T	12	85004551	C	T	0.0348	0.0348	0.0348	0.9870	0.9970	0.9340	0.9990
12_96027759_A_G	12	96027759	A	G	-0.0867	-0.0867	-0.0867	1.0000	1.0000	1.0000	1.0000
13_32839990_G_A	13	32839990	G	A	0.0424	0.0424	0.0424	1.0000	0.7550	1.0000	0.4080
13_32972626_A_T	13	32972626	A	T	0.2687	0.2308	0.4284	1.0000	1.0000	1.0000	1.0000
13_43501356_A_G	13	43501356	A	G	0.0517	0.0458	0.0975	0.9720	0.9430	0.9540	0.9500
13_73806982_T_C	13	73806982	T	C	0.0345	0.0251	0.0653	1.0000	0.9890	1.0000	0.9870
13_73960952_A_G	13	73960952	A	G	0.0399	0.0368	0.0730	1.0000	1.0000	1.0000	0.9980
14_105213978_T_G	14	105213978	T	G	0.0399	0.0399	0.0399	0.9830	0.9700	0.9440	0.9430
14_37128564_C_A	14	37128564	C	A	-0.0733	-0.0850	-0.0339	1.0000	0.9980	1.0000	0.9900
14_37228504_C_T	14	37228504	C	T	0.0390	0.0390	0.0390	0.9980	0.7800	0.9990	0.7530
14_68660428_T_C	14	68660428	T	C	-0.0474	-0.0612	0.0245	1.0000	1.0000	1.0000	1.0000
14_68979835_T_C	14	68979835	T	C	-0.0911	-0.0911	-0.0911	1.0000	1.0000	1.0000	1.0000
14_91751788_TC_T	14	91751788	TC	T	0.0380	0.0447	0.0091	1.0000	0.9810	1.0000	0.9860
14_91841069_A_G	14	91841069	A	G	0.0513	0.0513	0.0513	1.0000	1.0000	1.0000	1.0000
14_93070286_C_T	14	93070286	C	T	-0.0577	-0.0577	-0.0577	1.0000	0.9160	1.0000	0.7240
15_100905819_A_C	15	100905819	A	C	-0.0608	-0.0608	-0.0608	0.9520	0.7880	0.9400	0.6890
15_46680811_C_A	15	46680811	C	A	-0.1973	-0.1973	-0.1973	0.9070	0.7290	0.1420	0.2330
15_50694306_A_G	15	50694306	A	G	-0.0417	-0.0417	-0.0417	0.9600	0.8200	0.8510	0.7600
15_66630569_G_A	15	66630569	G	A	-0.0369	-0.0369	-0.0369	0.9770	0.9780	0.9690	0.9790
15_67457698_A_G	15	67457698	A	G	0.0782	0.0990	0.0141	1.0000	1.0000	1.0000	1.0000
15_75750383_T_C	15	75750383	T	C	-0.0413	-0.0413	-0.0413	0.9690	0.9740	0.9270	0.9390
15_91512267_G_T	15	91512267	G	T	-0.0589	-0.0589	-0.0589	1.0000	1.0000	1.0000	1.0000
16_10706580_G_A	16	10706580	G	A	-0.0740	-0.0740	-0.0740	0.9740	0.8340	0.9830	0.8520
16_23007047_G_T	16	23007047	G	T	0.1218	0.1218	0.1218	0.9070	0.7030	0.8410	0.5800
16_4008542_CAAAAA_C	16	4008542	CAAAAA	C	-0.0329	-0.0184	-0.0892	0.9080	0.7000	0.7390	0.5200
16_4106788_C_A	16	4106788	C	A	-0.0300	-0.0182	-0.0782	0.9740	0.7230	0.9680	0.5910
16_52538825_C_A	16	52538825	C	A	0.1147	0.1147	0.1147	0.9990	0.9990	0.9980	0.9980
16_52599188_C_T	16	52599188	C	T	0.1070	0.1070	0.1070	1.0000	0.9840	1.0000	0.9760
16_53809123_C_T	16	53809123	C	T	-0.0704	-0.0651	-0.0957	1.0000	0.9990	1.0000	0.9970
16_53861139_C_T	16	53861139	C	T	-0.0338	-0.0167	-0.0782	1.0000	1.0000	1.0000	0.9990
16_53861592_G_A	16	53861592	G	A	-0.0337	-0.0337	-0.0337	1.0000	0.9950	1.0000	0.9850
16_54682064_G_A	16	54682064	G	A	0.0477	0.0477	0.0477	1.0000	0.9400	1.0000	0.9660
16_6963972_C_G	16	6963972	C	G	0.0354	0.0303	0.0811	0.9650	0.8500	0.9360	0.6790
16_80648296_A_G	16	80648296	A	G	0.0839	0.0890	0.0467	1.0000	0.9980	1.0000	0.9840
16_85145977_T_C	16	85145977	T	C	-0.0211	-0.0044	-0.0714	0.9600	0.6480	0.9610	0.6880
16_87086492_T_C	16	87086492	T	C	-0.0469	-0.0469	-0.0469	0.9960	0.9530	0.9810	0.8880
17_29168077_G_T	17	29168077	G	T	-0.0568	-0.0568	-0.0568	0.9870	0.9830	0.9830	0.9790
17_39251123_T_C	17	39251123	T	C	0.0799	0.0631	0.1431	0.9090	0.8050	0.6200	0.5620
17_40127060_T_C	17	40127060	T	C	0.0174	-0.0161	0.1511	1.0000	0.5320	1.0000	0.2820
17_40485239_G_T	17	40485239	G	T	-0.0571	-0.0416	-0.1142	0.9470	0.9450	0.8060	0.6860
17_40744470_G_A	17	40744470	G	A	0.2017	0.2017	0.2017	0.9650	0.8910	0.7380	0.6510
17_43212339_C_CT	17	43212339	C	CT	0.0438	0.0438	0.0438	0.9710	0.8140	0.9300	0.6820
17_44283858_G_A	17	44283858	G	A	-0.0540	-0.0540	-0.0540	0.9580	0.9010	0.9350	0.9150
17_53209774_A_C	17	53209774	A	C	-0.0793	-0.0933	-0.0365	0.9980	0.9990	1.0000	1.0000
17_77781725_A_G	17	77781725	A	G	-0.0401	-0.0401	-0.0401	1.0000	0.9640	1.0000	0.9710
18_11696613_C_T	18	11696613	C	T	-0.0381	-0.0281	-0.0940	1.0000	0.6500	1.0000	0.5800
18_20634253_C_T	18	20634253	C	T	-0.0415	-0.0415	-0.0415	0.9810	0.9380	0.9920	0.9690
18_24125857_T_C	18	24125857	T	C	0.0346	0.0346	0.0346	1.0000	0.9680	1.0000	0.9150
18_24337424_C_G	18	24337424	C	G	0.0455	0.0455	0.0455	1.0000	1.0000	1.0000	1.0000
18_24518050_AT_A	18	24518050	AT	A	-0.0599	-0.0830	0.0060	1.0000	0.9390	1.0000	0.8880
18_25407513_C_G	18	25407513	C	G	0.0399	0.0307	0.0648	0.9800	0.9890	0.9650	0.9890
18_29981526_G_A	18	29981526	G	A	-0.1058	-0.1058	-0.1058	0.9990	0.9890	0.9970	0.9910
18_42411803_G_C	18	42411803	G	C	-0.0877	-0.1037	-0.0189	0.9980	0.9730	0.9980	0.9840
18_42888797_T_C	18	42888797	T	C	-0.0542	-0.0542	-0.0542	1.0000	1.0000	1.0000	1.0000

19_13249921_G_T	19	13249921	G	T	0.0956	0.0956	0.0956	0.9820	0.8510	0.7220	0.5750
19_17393925_C_A	19	17393925	C	A	0.0378	0.0036	0.1692	1.0000	1.0000	1.0000	0.9990
19_18569492_C_T	19	18569492	C	T	-0.0719	-0.0719	-0.0719	1.0000	0.9960	1.0000	0.9990
19_19517054_C_CGGGCG	19	19517054	C	CGGGCG	0.0437	0.0437	0.0437	0.9990	0.9470	0.9980	0.9440
19_44283031_T_C	19	44283031	T	C	0.0619	0.0619	0.0619	1.0000	0.8500	1.0000	0.8770
19_46166073_T_C	19	46166073	T	C	-0.0360	-0.0447	-0.0117	0.9500	0.8260	0.9430	0.7760
19_55816678_C_T	19	55816678	C	T	-0.0359	-0.0359	-0.0359	0.9910	0.9890	0.9860	0.9870
20_11379842_T_C	20	11379842	T	C	0.0844	0.0844	0.0844	0.9500	0.7680	0.9010	0.5330
20_41613706_C_G	20	41613706	C	G	0.0315	0.0266	0.0784	0.9490	0.9150	0.9260	0.8390
20_52296849_G_A	20	52296849	G	A	0.0440	0.0539	0.0144	0.9520	1.0000	0.9150	1.0000
20_5948227_G_A	20	5948227	G	A	0.0760	0.0760	0.0760	1.0000	1.0000	1.0000	1.0000
21_16364756_T_G	21	16364756	T	G	0.0646	0.0646	0.0646	1.0000	0.6250	1.0000	0.6540
21_16566350_A_G	21	16566350	A	G	0.0595	0.0678	0.0172	1.0000	0.9710	1.0000	0.8950
21_16574455_C_A	21	16574455	C	A	-0.0707	-0.0808	-0.0329	0.9980	0.9920	0.9940	0.9880
21_47762932_G_A	21	47762932	G	A	0.0946	0.0946	0.0946	0.9770	0.9370	0.9240	0.6980
22_19766137_C_T	22	19766137	C	T	-0.0367	-0.0367	-0.0367	0.9850	0.9370	0.9710	0.9130
22_29121087_A_G	22	29121087	A	G	0.1839	0.2812	-0.1566	1.0000	1.0000	1.0000	0.0030
22_29135543_G_A	22	29135543	G	A	0.0654	0.0654	0.0654	0.9970	0.9980	0.9950	0.9900
22_29203724_C_T	22	29203724	C	T	0.1405	0.1793	0.0191	1.0000	0.9480	0.9670	0.8150
22_29551872_A_G	22	29551872	A	G	-0.1716	-0.1716	-0.1716	0.9140	0.9080	0.8540	0.6280
22_38583315_AAAAG_AAAA	22	38583315	AAAAG	AAAAGAAAG	-0.0471	-0.0608	0.0079	0.9590	0.9360	0.9800	0.9640
22_39343916_T_A	22	39343916	T	A	0.0407	0.0407	0.0407	1.0000	0.8980	1.0000	0.7970
22_40904707_CT_C	22	40904707	CT	C	0.1148	0.1148	0.1148	0.9680	0.9630	0.9960	0.9830
22_43433100_C_T	22	43433100	C	T	-0.0600	-0.0600	-0.0600	0.9940	0.9960	0.9890	0.9920
22_45319953_G_A	22	45319953	G	A	-0.0134	-0.0060	-0.0611	0.9960	0.9200	0.9980	0.9200
22_46283297_G_A	22	46283297	G	A	0.0736	0.0736	0.0736	0.9730	0.6320	0.7520	0.4790

[a] Previously published by Mavaddat et al. (2019

[b] ER-specific PRS was constructed using a hybrid method, as described by Mavaddat et al. (201

[c] If r^2 =1 the variant was genotypec