

ORIGINAL ARTICLE

Analysis of rare disruptive germline mutations in 2135 enriched *BRCA*-negative breast cancers excludes additional high-impact susceptibility genes

C. Loveday¹, A. Garrett¹, P. Law¹, S. Hanks¹, E. Poyastro-Pearson¹, J. W. Adlard², J. Barwell³, J. Berg⁴, A. F. Brady⁵, C. Brewer⁶, C. Chapman⁷, J. Cook⁸, R. Davidson⁹, A. Donaldson¹⁰, F. Douglas¹¹, L. Greenhalgh¹², A. Henderson¹³, L. Izatt¹⁴, A. Kumar¹⁵, F. Lalloo¹⁶, Z. Miedzzybrodzka¹⁷, P. J. Morrison¹⁸, J. Paterson¹⁹, M. Porteous²⁰, M. T. Rogers²¹, L. Walker²², The Breast and Ovarian Cancer Susceptibility Collaboration[†], D. Eccles²³, D.G. Evans¹⁶, K. Snape^{1,24}, H. Hanson^{1,24}, R.S. Houlston¹ & C. Turnbull^{1,25*}

¹Division of Genetics and Epidemiology, Institute of Cancer Research, London; ²Yorkshire Regional Genetics Service, St James's University Hospital, Leeds; ³Leicestershire Genetics Centre, University Hospitals of Leicester National Health Service (NHS) Trust, Leicester; ⁴Division of Medical Sciences, Human Genetics, University of Dundee, Dundee; ⁵North West Thames Regional Genetics Service, Kennedy Galton Centre, London; ⁶Peninsula Regional Genetics Service, Royal Devon & Exeter Hospital, Exeter; ⁷West Midlands Regional Genetics Service, Birmingham Women's Hospital, Birmingham; ⁸Sheffield Regional Genetics Service, Sheffield Children's NHS Foundation Trust, Sheffield; ⁹West of Scotland Regional Genetics Service, Ferguson Smith Centre for Clinical Genetics, Glasgow; ¹⁰South Western Regional Genetics Service, University Hospitals of Bristol NHS Foundation Trust, Bristol; ¹¹Northern Genetics Service, Newcastle upon Tyne Hospitals NHS Foundation Trust, Newcastle upon Tyne; ¹²Cheshire and Merseyside Clinical Genetics Service, Alder Hey Children's NHS Foundation Trust, Liverpool; ¹³Northern Genetics Service (Cumbria), Newcastle upon Tyne Hospitals NHS Trust, Newcastle upon Tyne; ¹⁴South East Thames Regional Genetics Service, Guy's and St. Thomas NHS Foundation Trust, London; ¹⁵North East Thames Regional Genetics Service, Great Ormond St. Hospital, London; ¹⁶University Department of Medical Genetics & Regional Genetics Service, St. Mary's Hospital, Manchester; ¹⁷University of Aberdeen and North of Scotland Clinical Genetics Service, Aberdeen Royal Infirmary, Aberdeen; ¹⁸Belfast Health and Social Care (HSC) Trust & Department of Medical Genetics, Northern Ireland Regional Genetics Service, Queen's University Belfast, Belfast; ¹⁹East Anglian Regional Genetics Service, Cambridge University Hospitals NHS Foundation Trust, Cambridge; ²⁰South East of Scotland Clinical Genetics Service, Western General Hospital, Edinburgh; ²¹All Wales Medical Genetics Service, University Hospital of Wales, Cardiff; ²²Oxford Regional Genetics Service, Oxford Radcliffe Hospitals NHS Trust, Oxford; ²³Faculty of Medicine, University of Southampton, Southampton University Hospitals NHS Trust, Southampton; ²⁴South West Thames Regional Genetics Service, St. George's Hospital, London; ²⁵Royal Marsden NHS Foundation Hospital, London, UK



Available online 17 September 2022

Background: Breast cancer has a significant heritable basis, of which ~60% remains unexplained. Testing for *BRCA1/BRCA2* offers useful discrimination of breast cancer risk within families, and identification of additional breast cancer susceptibility genes could offer clinical utility.

Patients and methods: We included 2135 invasive breast cancer cases recruited via the Breast and Ovarian Cancer Susceptibility study, a retrospective UK study of familial breast cancer. Eligibility criteria: female, *BRCA*-negative, white European ethnicity, and one of: (i) breast cancer family history, (ii) bilateral disease, (iii) young age of onset (<30 years), and (iv) concomitant ovarian cancer.

We undertook exome sequencing of cases and carried out gene-level burden testing of rare damaging variants against those from 51 377 ethnicity-matched population controls from gnomAD.

Results: 159/2135 (7.4%) cases had a qualifying variant in an established breast cancer susceptibility gene, with minimal evidence of signal in other cancer susceptibility genes. Known breast cancer susceptibility genes *PALB2*, *CHEK2*, and *ATM* were the only genes to retain statistical significance after correcting for multiple testing. Due to the enrichment of hereditary cases in the series, we had good power (>80%) to detect a gene of *BRCA1*-like risk [odds ratio (OR) = 10.6] down to a population minor allele frequency of 4.6×10^{-5} (1 in 10 799, less than one-tenth that of *BRCA1*) and of *PALB2*-like risk (OR = 5.0) down to a population minor allele frequency of 2.8×10^{-4} (1 in 1779, less than half that of *PALB2*). Power was lower for identification of novel moderate penetrance genes (OR = 2-3) like *CHEK2* and *ATM*.

Conclusions: This is the largest case-control whole-exome analysis of enriched breast cancer published to date. Whilst additional breast cancer susceptibility genes likely exist, those of high penetrance are likely to be of very low mutational frequency. Contention exists regarding the clinical utility of such genes.

*Correspondence to: Prof. Clare Turnbull, Division of Genetics and Epidemiology, Institute of Cancer Research, 15 Cotswold Road, Sutton, Surrey, SM2 5NG, UK. Tel: +44-208-722-4485

E-mail: clare.turnbull@icr.ac.uk (C. Turnbull).

Manuscript social media information: [@clare_turnbull](https://twitter.com/clare_turnbull), [@LovedayChey](https://twitter.com/LovedayChey), [@DrAliceGarrett](https://twitter.com/DrAliceGarrett), [@ICR_London](https://twitter.com/ICR_London), www.icr.ac.uk.

[†]Individual collaborators and their affiliations are listed in the Appendix. 0923-7534/© 2022 The Author(s). Published by Elsevier Ltd on behalf of European Society for Medical Oncology. This is an open access article under the CC BY license (<http://creativecommons.org/licenses/by/4.0/>).

Key words: breast cancer, genetic susceptibility, cancer susceptibility genes, whole-exome sequencing, rare-variant burden testing

INTRODUCTION

Female breast cancer is now the most common cancer with 2.3 million cases diagnosed annually worldwide and a lifetime risk of ~15%.^{1,2} The heritable basis of the disease is well established and evidenced by the increased risk in relatives of cases.^{3,4} It is now well recognized that the genetic architecture of breast cancer susceptibility encompasses a broad spectrum, from common polymorphisms individually conferring small risks through to breast cancer susceptibility genes (BCSGs; [Supplementary Table S1](https://doi.org/10.1016/j.annonc.2022.09.152), available at <https://doi.org/10.1016/j.annonc.2022.09.152>) characterized by multiple disparate mutations. BCSGs are typically categorized by their disease penetrance, for example, as very high penetrance [odds ratio (OR) ≥ 10], high (OR ≥ 5), or moderate/intermediate (OR ≥ 2).⁵

It is estimated that ~41% of the heritable risk of breast cancer has been deciphered.⁶ Of the 23% attributable to rare variants, the majority is accounted for by *BRCA1* and *BRCA2* (17%), with other established BCSGs (*PALB2*, *ATM*, *CHEK2*, *TP53*, *STK11*, *PTEN*, *NF1*, *CDH1*, *BARD1*, *RAD51C*, *RAD51D*) collectively accounting for about 6%.⁷ It is estimated that 18% of genetic risk of breast cancer resides in common variants identified by genome-wide association studies (GWAS), with statistical modelling suggesting a further 23% is potentially tractable by larger GWAS.⁸ Of the 59% of unexplained heritable risk, over half (36%) is thus potentially ascribable to hitherto unidentified rare variants. Whilst many of these variants may be non-coding and/or of low penetrance, there is considerable interest in the component that may be enshrined in further susceptibility genes of moderate to high penetrance which are both mechanistically important as well as potentially relevant to the clinical management of patients and their families. Numerous studies have used next-generation sequencing (NGS) technologies to search this space for novel BCSGs ([Supplementary Table S2](https://doi.org/10.1016/j.annonc.2022.09.152), available at <https://doi.org/10.1016/j.annonc.2022.09.152>), though few if any convincing candidates to date have been identified.

To look for novel BCSGs, we carried out germline whole-exome sequencing (WES) of 2135 *BRCA1/BRCA2*-negative genetically enriched breast cancer cases. We used a three-phase analytical strategy ([Figure 1](#)). Firstly, we determined the contribution of pathogenic variants in established autosomal dominant cancer susceptibility genes (CSGs; $n = 86$ genes). Secondly, we assessed the contribution of candidate genes implicated on the basis of involvement in DNA repair ($n = 276$ genes), and/or a known oncogenomic role in cancer ($n = 686$ genes) by carrying out burden analyses collapsing at the gene level. Finally, we carried out exome-wide rare-variant burden testing, undertaking correction for cumulative performance of multiple tests.

PATIENTS AND METHODS

Subjects and data

We included *BRCA1/BRCA2*-negative women with a diagnosis of pathologically confirmed invasive breast cancer ($n = 2430$), recruited to the Institute of Cancer Research Breast and Ovarian Cancer Susceptibility (BOCS) study. Eligibility criteria included one of: (i) a significant family history of breast cancer (breast cancer in at least one first-degree relative or two second-degree relatives); (ii) bilateral breast cancer; (iii) early-onset breast cancer (<30 years); and (iv) concomitant ovarian cancer.⁹ Cases were assigned a 'genetic enrichment score' based on allele sharing to reflect the degree of personal and family history of breast cancer as follows: bilateral disease in the proband received a score of 1, with additional points assigned for disease in first-degree relatives (0.5 each), second-degree relatives (0.25 each), and third-degree relatives (0.125 each). We used a score of ≥ 1.5 to define a set of 'high-risk' cases ($n = 855$).

The proportion of cases of self-reported Ashkenazi Jewish ancestry was <1%.

Written informed consent was obtained from all participants and the research was approved by the London Multicentre Research Ethics Committee (MREC/01/2/18).

For controls, we used WES data from 51 377 non-Finnish European non-cancer individuals from the Genome Aggregation Database (gnomAD) (date of accession 21 May 2020).

Whole-exome sequencing and quality control

Germline DNA from cases underwent WES using Illumina technology. Data processing, variant calling, and annotation were carried out according to Genome Analysis Toolkit (GATK) best practices. We excluded cases based on quality, cryptic familial duplicates, presence of occult *BRCA1/BRCA2* mutations, and ancestry analysis using principal components ([Supplementary Figure S1](#), available at <https://doi.org/10.1016/j.annonc.2022.09.152>, [Supplementary Table S3](#), available at <https://doi.org/10.1016/j.annonc.2022.09.152>).

Case-control variant set alignment

Variant-level exclusions are described in [Supplementary Table S4](#), available at <https://doi.org/10.1016/j.annonc.2022.09.152> and [Supplementary methods](#), available at <https://doi.org/10.1016/j.annonc.2022.09.152>. Analyses were restricted to rare variants with a minor allele frequency (MAF) < 0.5%. We controlled against spurious inflation of the burden test statistic by filtering the bottom fifth centile of variants based on the GATK variant quality metric QualByDepth (QD) and restricting analyses to base positions with comparable levels of

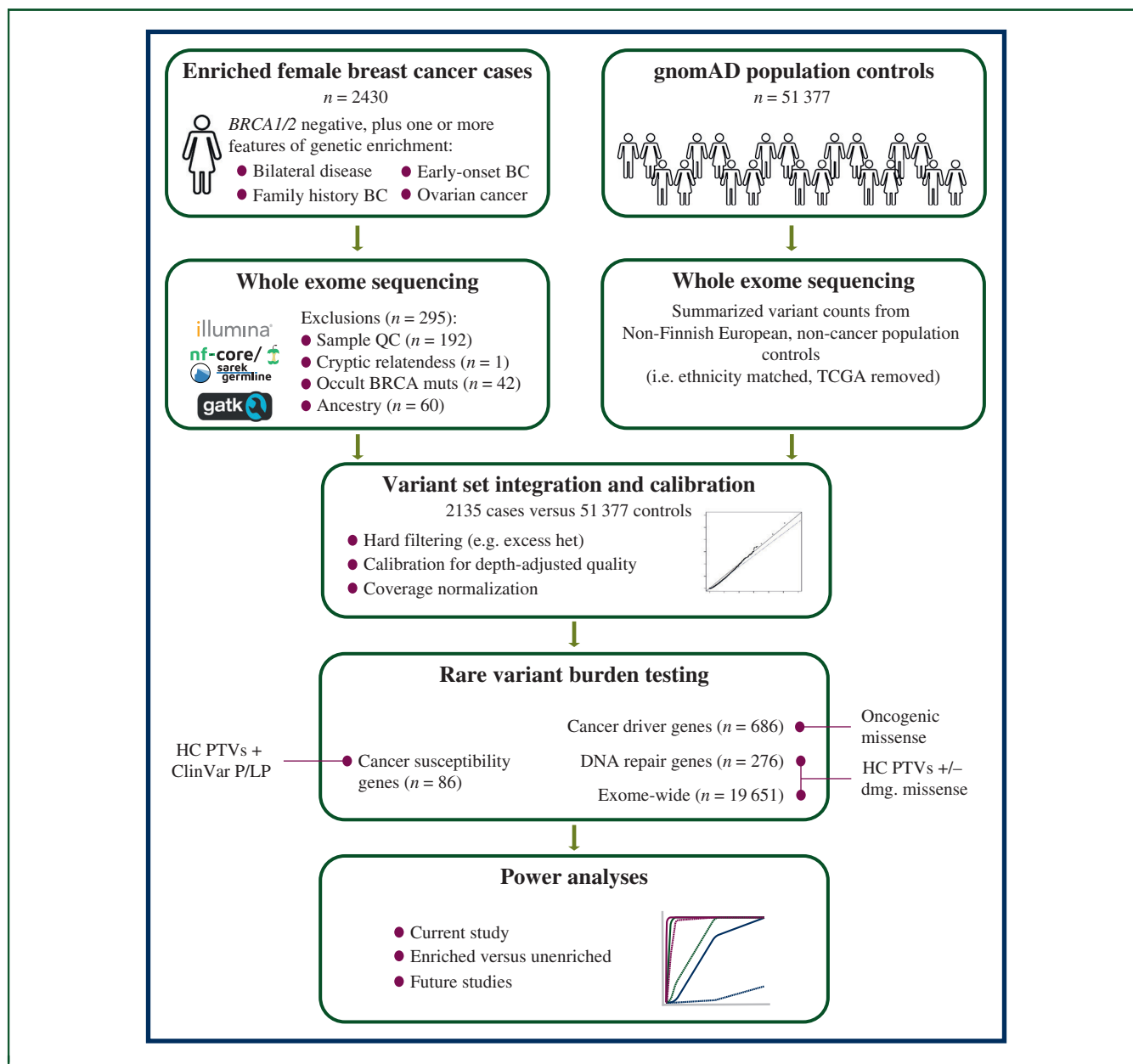


Figure 1. Study Overview. Overview of case and control samples used, along with generation and integration of datasets and analyses carried out. BC, breast cancer; dmg. missense, damaging missense variant; HC-PTV, high-confidence protein truncating variant; QC, quality control; TCGA, The Cancer Genome Atlas.

coverage (Supplementary Figures S2 and S3, available at <https://doi.org/10.1016/j.annonc.2022.09.152>, Supplementary Table S5, available at <https://doi.org/10.1016/j.annonc.2022.09.152>, Supplementary methods, available at <https://doi.org/10.1016/j.annonc.2022.09.152>).¹⁰

Statistical analyses

We carried out collapsed gene-level burden analyses in: (i) autosomal dominantly acting CSGs ($n = 86$)^{11,12}; (ii) DNA repair genes ($n = 276$)¹³; (iii) somatic cancer driver genes ($n = 686$ genes)¹⁴; and (iv) the exome ($n = 19\,651$ genes) (Supplementary Table S6, available at <https://doi.org/10.1016/j.annonc.2022.09.152>).

A single class of test was carried out for the CSGs, comprising high-confidence protein truncating variants (HC-PTVs, as defined by Loss-of-Function Transcript Effect Estimator)¹⁵ plus variants (regardless of variant consequence or LOFTEE annotation) classified as pathogenic/likely pathogenic (P/LP) on ClinVar with one or more stars using The American College of Medical Genetics and Genomics criteria.¹⁶ A single class of test was also carried out for missense/in-frame variants at recurrently somatically mutated residues in cancer driver genes (as defined in oncoKB; Supplementary Table S7, available at <https://doi.org/10.1016/j.annonc.2022.09.152>). For the DNA repair/exome-wide analyses, two classes of test were carried out: class I (HC-PTVs only) and class II (HC-PTVs + missense

with REVEL score > 0.7).^{17,18} The proportion of REVEL-qualifying missense variants that were also pathogenic in ClinVar in known CSGs was 9/172 (5.2%).

Association was measured using Fisher's exact test (two-sided) comparing case-control proportions with a qualifying variant. For declaring statistical significance, we imposed a Bonferroni correction of $P < 0.05$ for the number tests carried out, accounting for cumulative testing in candidate-based followed by exome-wide analyses. For replication, we used WES data from UK Biobank comprising 15 017 female breast cancer cases and 199 479 healthy women with no previous history of cancer, assigning $P < 0.05$ as nominal evidence of replication.¹⁹ Analyses were carried out in R (v4.0.3).

Calculation of study power

Disease allele frequency in controls was taken as the baseline, and the frequency in cases was determined by a weighted average of the predicted enrichment in cases with bilateral disease and/or an affected first-degree relative.²⁰ Fisher's exact test (two-sided) was carried out for each sampling of cases and controls, carried out 10 000 times for each frequency/relative risk combination.

Additional methodological information is available in the [Supplementary methods](https://doi.org/10.1016/j.annonc.2022.09.152), available at <https://doi.org/10.1016/j.annonc.2022.09.152>.

RESULTS

The final case dataset comprised 2135 *BRCA*-negative breast cancer cases, of which 838 (39%) had bilateral disease, 1628 (76%) had a significant family history of breast cancer (at least one first-degree relative or two second-degree relatives), 183 (9%) had breast cancer diagnosed ≤ 30 years, and 261 (12%) had concomitant ovarian cancer ([Figure 2](#)). The cases were enriched for bilaterality, family history, and early-onset disease compared to unselected cases from the literature

([Supplementary Figure S4](#), available at <https://doi.org/10.1016/j.annonc.2022.09.152>).^{21,22} Mean age at first diagnosis was 48 years (compared to the UK average of >60 years¹).

Frequency of pathogenic germline variants in cancer susceptibility genes

A total of 159/2135 (7.4%) of cases had a qualifying variant (HC-PTV and/or ClinVar P/LP) in one of the 13 established BCSGs, including 44 (2.1%) in *PALB2*, 41 (1.9%) in *ATM*, and 53 (2.5%) in *CHEK2* ([Figure 3](#), [Supplementary Table S8](#), available at <https://doi.org/10.1016/j.annonc.2022.09.152>). Eleven cases (0.5%) had a variant in genes associated with estrogen receptor (ER)-negative disease: *BARD1* (5, 0.2%), *RAD51C* (4, 0.2%), *RAD51D* (2, 0.1%). Eleven cases (0.5%) had a variant in a rare pleiomorphic cancer susceptibility gene, comprising *TP53* (7, 0.3%), *CDH1* (1, 0.05%), *NF1* (2, 0.1%), and *PTEN* (1, 0.05%). There was no significant difference in the age of breast cancer diagnosis, burden of family history, or frequency of bilateral disease between cases with and without a qualifying variant ($P > 0.05$ in relevant statistical test: see [Supplementary methods](#), available at <https://doi.org/10.1016/j.annonc.2022.09.152>).

There were no significant associations with breast cancer amongst the other CSGs, though there was an appreciable case frequency of qualifying variants in some, notably *MSH6* (10 cases) and *BRIP1* (5 cases), genes previously proposed as BCSGs ([Figure 3](#), [Supplementary Table S8](#), available at <https://doi.org/10.1016/j.annonc.2022.09.152>).^{23,24} Excluding the 13 established BCSGs, the frequency of qualifying variants in CSGs was not significantly elevated in cases (53/2135 or 2.5%) compared to controls (1365/51 377 or 2.7%) ($P = 0.68$).

Candidate gene burden testing

Cognizant of the impact of correction for multiple testing, we firstly considered targeted candidacy-based approaches. On the functional basis of genes implicated to date in breast

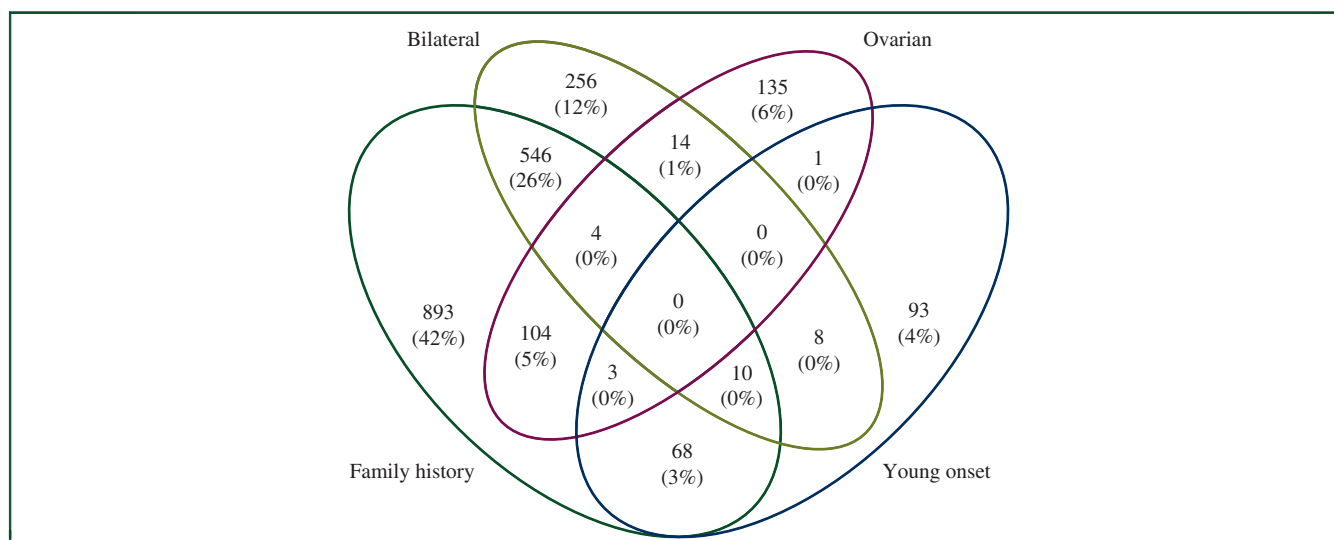


Figure 2. Distribution of case phenotypic features. Distribution of phenotypes across 2135 probands with invasive breast cancer where 'Family history' denotes at least one first-degree or two second-degree relatives affected with breast cancer, 'Ovarian' denotes the proband also being diagnosed with invasive ovarian cancer, 'Young onset' denotes the proband's first cancer being diagnosed at age <30 , and 'Bilateral' denotes a diagnosis of bilateral breast cancer.

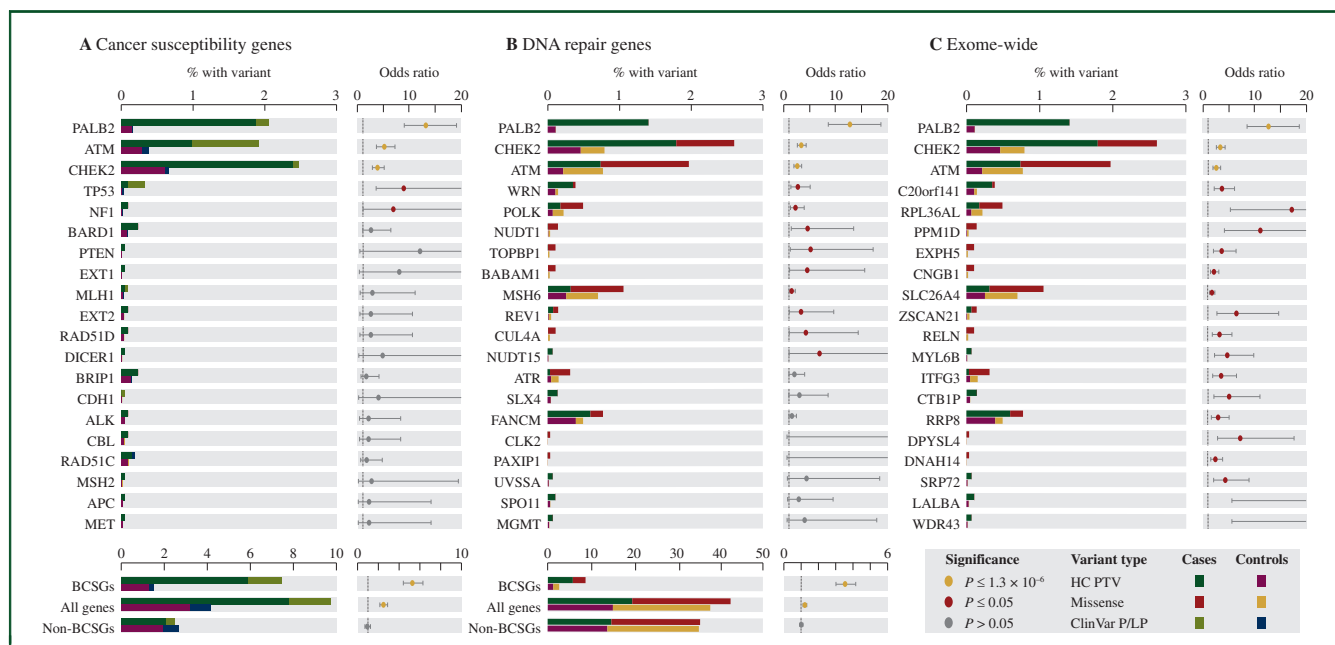


Figure 3. Frequency of deleterious variants in cases versus controls. Frequency of rare variants (MAF < 0.5%), collapsed by gene, comparing cases to controls for (A) high-confidence protein truncating variants (HC-PTVs, as defined by LOFTEE) plus pathogenic/likely pathogenic variants (defined by ≥ 1 star review status ClinVar entries) in 86 established autosomal dominant cancer susceptibility genes; (B) HC-PTVs plus damaging missense with REVEL score ≥ 0.7 in 276 genes involved in DNA repair; (C) HC-PTVs plus damaging missense exome-wide. The left panel in each sub-figure shows the percentage of cases/controls and the right panel shows the corresponding odds ratio. Genes are ordered by P value. For each analysis, only the top 20 associations have been shown and the comparison is also shown for all genes, all breast cancer susceptibility genes (BCSGs), and all non-BCSGs.

cancer susceptibility, we first considered 276 genes involved in DNA repair and carried out burden analysis for rare variants (MAF < 0.5%) of class I (HC-PTVs only) and class II (HC-PTVs and deleterious non-synonymous variants). *PALB2*, *CHEK2*, and *ATM* emerged from this analysis as strongly associated (Figure 3, Supplementary Table S8, available at <https://doi.org/10.1016/j.annonc.2022.09.152>). *WRN* was the highest ranked (non-BCSG) DNA repair gene by P value, driven particularly by the HC-PTV frequency, but the association did not retain significance following Bonferroni correction (class I test: 10/2135, 0.47% cases versus 73/51 377, 0.14% controls, $P = 0.002$). There was only nominal statistical support on querying of the UK Biobank data for association of this gene with breast cancer, ($P = 0.0087$), which again did not withstand correction for multiple testing (Supplementary Table S9, available at <https://doi.org/10.1016/j.annonc.2022.09.152>). Excluding the known BCSGs, collapsing across the remaining 263 genes involved in DNA repair the signal of association was not significant on analysis of variants of class I (316/2135 versus 7130/51377, $P = 0.23$) or class II (756/2135 versus 18 048/51 377, $P = 0.8$).

Next, on the basis of exemplar genes associated with both autosomal dominant cancer susceptibility and somatic oncogenic mechanisms [e.g. *RET* mutations predisposing to thyroid cancer and *KIT* mutations predisposing to gastrointestinal stromal tumor (GIST)], we considered 686 cancer driver genes for which recurrent somatic hotspot mutations have been reported. On burden analysis for rare missense/in-frame variants (MAF < 0.5%) overlapping with known oncogenic residues (Supplementary Figure S5, available at <https://doi.org/10.1016/j.annonc.2022.09.152>, Supplementary Table S8,

available at <https://doi.org/10.1016/j.annonc.2022.09.152>), there were no significant associations after correcting for multiple testing.

Exome-wide rare-variant burden analysis

We carried out exome-wide burden testing, again collapsing at the gene level for association of rare variants (MAF < 0.5%) of each of class I and class II. Again, only for *PALB2* (classes I and II), *CHEK2* (classes I and II), and *ATM* (class II only) were the associations significant following Bonferroni correction (Figure 3, Supplementary Table S8, available at <https://doi.org/10.1016/j.annonc.2022.09.152>). For all three genes, the ORs for breast cancer risk with class I variants were substantially inflated compared to published estimates from unselected breast cancers,²⁵ commensurate with the inflation predicted on the basis of the genetic enrichment in our cases (Supplementary Table S8, available at <https://doi.org/10.1016/j.annonc.2022.09.152>). Amongst the top 20 associations ranked by the smallest P value were *C20orf141* (class I and class II) and *EXPH5* (class I and class II), but the associations were not significant following Bonferroni correction. Furthermore, neither these genes nor any others in the top 20 associations showed evidence of replication in the UK Biobank ($P > 0.1$; Supplementary Table S9, available at <https://doi.org/10.1016/j.annonc.2022.09.152>).

Notable amongst the top 20 ranked gene associations in this analysis was *PPM1D*. Erroneously reported originally as a BCSG, *PPM1D* is characterized by somatic mosaic mutations arising in hematopoietic cells in response to chemotherapy.²⁶⁻²⁸ Consistent with somatic mosaicism in blood

cells, the *PPM1D* VAFs detected in this study were skewed downwards for PTVs compared to synonymous variants (Kolmogorov–Smirnov P value = 0.11, Mood's median P value = 0.17; [Supplementary Figure S6](https://doi.org/10.1016/j.annonc.2022.09.152), available at <https://doi.org/10.1016/j.annonc.2022.09.152>).

There was likewise no evidence of association for any genes on subsequent restriction to a smaller set of 855 'high-risk' cases (see Methods, [Supplementary Table S8](https://doi.org/10.1016/j.annonc.2022.09.152), available at <https://doi.org/10.1016/j.annonc.2022.09.152>).

On gene-set enrichment analysis of gene ontology terms and Kyoto Encyclopedia of Genes and Genomes pathways amongst genes with a class I/class II test P value ≤ 0.01 ,^{29,30} multiple gene groups showed significant associations, primarily related to DNA replication and repair, but once corrected for the signal of association for constituent known BCSGs no gene groups showed significant residual association ([Supplementary Table S10](https://doi.org/10.1016/j.annonc.2022.09.152), available at <https://doi.org/10.1016/j.annonc.2022.09.152>).

Power analyses

Recent large-scale genetic epidemiologic analyses have confirmed *BRCA1* as being of almost twice the penetrance but much lower mutational frequency than *BRCA2* (*BRCA1*: OR = 10.6, population $MAF_{combined}$ = 0.10% or 1 in 494; *BRCA2*: OR = 5.9, population $MAF_{combined}$ = 0.16% or 1 in 306).²⁵ However, *PALB2* (OR = 5.0, population $MAF_{combined}$ = 0.07% or 1 in 667) has substantially lower 'impact profile' (penetrance \times mutational frequency) with risk in the rough range of *BRCA2* but a mutational frequency similar to *BRCA1*. Taking into account the enrichment as well as the size of our case and control series, we had very good power (90%) to identify a new high-impact BCSG conferring a *BRCA1*-like breast cancer risk (OR = 10.6) with population-combined mutation frequency ($MAF_{combined}$) down to (i) 0.0043% (or 1 in 11 628) in the targeted DNA repair gene analysis or (ii) $MAF_{combined}$ of 0.0056% (or 1 in 8945) in the WES analysis. For a new gene of *PALB2*-like (OR = 5) or *ATM*-like (OR = 2.1) breast cancer risk, we had 90% power to identify a gene of $MAF_{combined}$ down to (i) 0.026% (1 in 1946) or 0.37% (1 in 135), respectively, in the DNA repair analysis and (ii) 0.034% (1 in 1475) and 0.5% (1 in 100) in the WES analysis ([Figure 4](https://doi.org/10.1016/j.annonc.2022.09.152), [Tables 1 and 2](https://doi.org/10.1016/j.annonc.2022.09.152)).

These estimates are likely to be inherently conservative, for although we have accounted for bilateral disease and the first affected first degree relative in our series where present, we have not accounted for the additional family members with breast cancer, young age of cases, and concomitant occurrence of ovarian cancer.²⁰

Power calculations representing variability in population $MAF_{combined}$, the largest breast WES study published to date, and expansion of the current study with increasing sample size, are presented in [Supplementary Tables S11, S12, and S13](https://doi.org/10.1016/j.annonc.2022.09.152), available at <https://doi.org/10.1016/j.annonc.2022.09.152>, respectively.

DISCUSSION

In our series of 2135 enriched *BRCA*-negative female breast cancer cases, only 159 (7.4%) have a P/LP variant in an

established BCSG. The vast majority (138/159, 87%) of these P/LP variants were in three genes: *PALB2*, *CHEK2*, and *ATM*. As would be anticipated from our ascertainment framework, the contribution of P/LP variants in genes associated with pleiomorphic tumor syndromes was very modest (0.5% of cases). Furthermore, given that P/LP variants in *BARD1*, *RAD51C*, and *RAD51D* are of relatively low frequency even in the triple-negative breast cancers with which they are associated, their contribution was predictably low in this series ascertained agnostic to histology (0.5% of cases). The frequency in cases of P/LP variants in other autosomal dominant CSGs (2.5%) was not elevated compared to controls (2.7%), suggesting firstly that these variants are largely incidental to etiology of these breast cancers and secondly that amongst other established CSGs there is minimal unexplained residual pleiotropy for breast cancer.

In our subsequent exploration of residual breast cancer genetic susceptibility attributable to rare variants, aside from 'rediscovery' of established BCSGs *PALB2*, *CHEK2*, and *ATM*, no gene attained statistically significant signal of association at the respective thresholds for either the candidate-based approaches or the exome-wide approach. Furthermore, there was no evidence for association found in a replication series (UK Biobank) for top-ranking genes with nonsignificant signals of association.

To our knowledge, our study is the largest case-control analysis reported to date in enriched breast cancer cases, and better powered for discovery of a new rare high-penetrance susceptibility gene than the largest reported case-control analysis to date of unselected breast cancers (7859 cases/117 456 controls)¹⁹ ([Supplementary Tables S2 and S12](https://doi.org/10.1016/j.annonc.2022.09.152), available at <https://doi.org/10.1016/j.annonc.2022.09.152>).

Contextualization of our findings with previous studies

In the 1990s, linkage analysis in modest numbers of multi-case pedigrees enabled identification of *BRCA1* followed by *BRCA2*.^{31–34} Subsequent linkage analyses in larger series of *BRCA1*- and *BRCA2*-negative breast cancer families yielded no reproducible signals.³⁵ Thereafter followed case-control mutational screening for genes functionally related to *BRCA1* and *BRCA2*, necessarily limited to modest numbers of samples on account of the low-throughput technologies available. Through these experiments, *PALB2*, *CHEK2*, and *ATM* were identified as BCSGs, with the less reproducible associations for *BARD1*, *RAD51C*, and *RAD51D* subsequently confirmed as exclusive to ER-negative tumors.^{25,36–38}

It was widely anticipated that the advent of NGS would allow identification of the elusive additional BCSGs. Initial applications of WES typically involved case-only segregation-type analyses of multiple members of breast cancer families ([Supplementary Table S2](https://doi.org/10.1016/j.annonc.2022.09.152), available at <https://doi.org/10.1016/j.annonc.2022.09.152>). With wider public availability of control data, NGS breast cancer gene discovery studies progressed on to case-control analyses, with the vast majority focused on targeted sets of DNA repair genes selected on the basis of biological candidacy.^{39–41} Few case-control WES

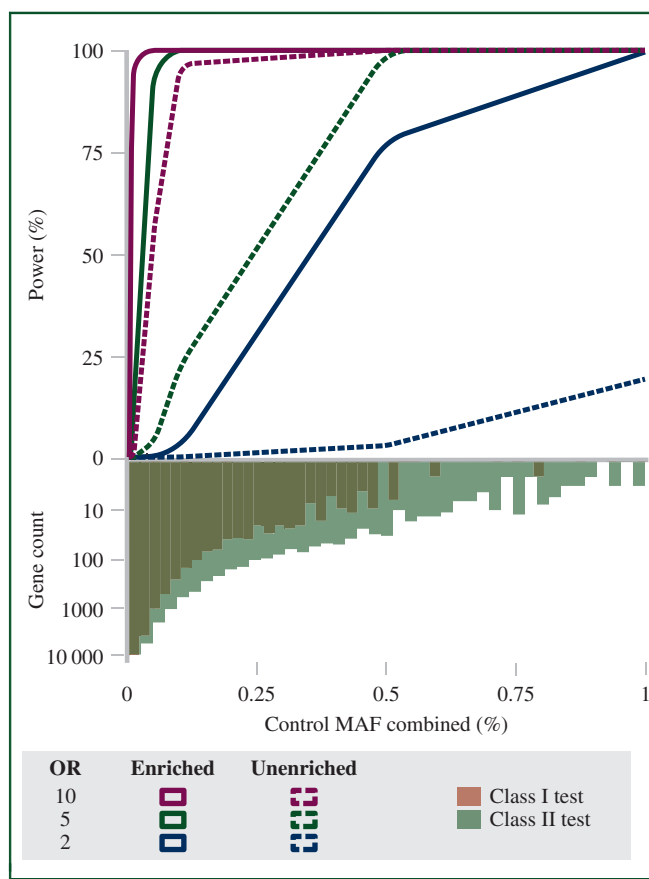


Figure 4. Power calculations in enriched versus unenriched breast cancer cases. The line plot demonstrates the power afforded by the current study of 2135 genetically enriched breast cancer cases (solid lines) versus an unenriched (unselected) series of the same size (dashed lines) at exome-wide significance ($P < 1.3 \times 10^{-06}$). The histogram plot shows the number of genes with the corresponding control MAF_{combined} for each of the two variant class tests. OR, odds ratio.

analyses for breast cancer have been reported. In 2021, Wang et al. reported the largest WES case-control analysis in breast cancer to date comprising 7859 unselected breast cancer cases compared to 117 456 controls in UK Biobank.¹⁹ Overall, despite almost a decade of application of NGS via family-only, candidate, and exome-wide case-control approaches, including the study presented here, none has yielded robust novel associations of genes with breast cancer susceptibility.

Power analyses from this study demonstrate that whilst entirely plausible that additional high-penetrance BCSGs exist, (i) for genes of *BRCA1*-like risks, the MAF_{combined} will be vanishingly low, and (ii) for any gene in the ~5-fold range of risk, the population mutational frequency is likely to be less than half that of *PALB2* (i.e. <1 in 1500 in the population). These sobering power analyses contextualize the failure of those modest-sized early linkage studies to identify *PALB2* and of the WES familial segregation studies to yield reproducible findings.

Conversely, our study was substantially less well powered for genes of more modest effect sizes, for which the use of enriched cases also offers less benefit (Tables 1 and 2). For truncating variants in *CHEK2*, large genetic epidemiologic studies have demonstrated OR of 2.5 and population

MAF_{combined} of 0.31% (1 in 163).²⁵ Our power to detect a gene of this risk/MAF_{combined} profile was 95%, consistent with *CHEK2* coming out in our analysis as statistically significant for class I. However, for *ATM* we only had 12% power for detection of class I variants as the penetrance is lower (OR = 2.1) and the population mutational frequency is half that of *CHEK2* [MAF_{combined} = 0.14% (1 in 350)]. Thus, it is entirely plausible that numerous additional moderate penetrance BCSGs exist of impact profiles (penetrance \times mutational frequency) akin to that of *ATM* and up to that of *CHEK2*, for which discovery studies to date (including ours) have been underpowered and/or not included in previously studied gene sets.

High-penetrance genes are widely agreed to offer high clinical utility. Even recognizing variation in background polygenic contribution, the breast cancer risks are deemed sufficiently high to dichotomize management within families, with aggressive interventions such as risk-reducing mastectomies justified in family members carrying the mutation and restoration to near-population risk for those without. Contention persists regarding the utility for risk stratification offered by genes of more moderate penetrance, especially when conferring only sub-type-specific risks and/or lacking recurrent founder mutations.⁵

Limitations of the study

Data regarding receptor status and/or histopathology were only available for a subset of patients, thus limiting power for histologically driven subtype analyses. The gnomAD series is only available as summary-level count data; as such, equivalently stringent sample- and variant-level quality control for the case data were required to address the possibility of inflation due to systematic genotyping error and potential differences in depth of coverage. This also meant we were unable to carry out robust analyses for recessive and co-dominant susceptibility.

Another limitation was definitions by which missense variants were included as pathogenic/deleterious. For the CSG analysis, we included only missense variants classified on ClinVar as pathogenic/likely pathogenic with at least one star. Thus, any variants assigned on ClinVar as having 'conflicting interpretations' were excluded. Likewise, for the DNA repair/exome-wide analyses, application of a threshold of REVEL > 0.7 was inevitably an imperfect filter, likely resulting both in inclusion of neutral variants and exclusion of deleterious variants.

Additionally, we did not include exon-level (or greater) deletions and duplications in our analyses. Such variants (which are inherently more difficult to detect from short-read NGS data) are thought to comprise a non-trivial fraction of disease-associated germline alterations in cancer susceptibility genes. In the case of *BRCA1* and *BRCA2*, they are thought to account for somewhere between 10% and 20% of all alterations in those genes. Thus, any putative novel breast cancer gene where structural variation is the predominant mutation type would be missed by our analyses. Our power analyses, therefore, do not account for any

Table 1. Power to discover genes of effect size and population frequency equivalent to known breast cancer susceptibility genes

Gene effect size	Published population frequency of gene (PTVs)		Power to detect (%) in study of 2135 cases versus 51 377 controls			
	Population MAF	Population frequency	Enriched		Un-enriched	
			DNA repair analysis	WES analysis	DNA repair analysis	WES analysis
<i>BRCA1</i> -like (OR = 10.6)	1.01E-03	1 in 494	100.0%	100.0%	99.8%	98.2%
<i>BRCA2</i> -like (OR = 5.9)	1.63E-03	1 in 306	100.0%	100.0%	94.0%	80.9%
<i>PALB2</i> -like (OR = 5)	7.49E-04	1 in 667	100.0%	100.0%	31.0%	12.2%
<i>CHEK2</i> -like (OR = 2.5)	3.07E-03	1 in 163	99.2%	95.6%	22.0%	6.1%
<i>ATM</i> -like (OR = 2.1)	1.43E-03	1 in 350	31.1%	11.5%	1.6%	0.2%

Power afforded by study of 2135 cases and 51 377 controls to discover genes of the OR/MAF profile of *BRCA1*, *BRCA2*, *PALB2*, *CHEK2*, and *ATM* with Bonferroni correction undertaking targeted ($n = 286$ DNA repair genes) or exome-wide analysis ($n = 19\ 651$ genes), comparing use of enriched versus unenriched cases. MAF, minor allele frequency; OR, odds ratio; PTVs, protein truncating variants; WES, whole-exome sequencing.

gain in power that would be afforded by the inclusion of such variants.

Future studies

Future analysis of 15 000 similarly enriched breast cancer cases against 400 000 ethnicity-matched controls would be fully powered (>99%) to identify all genes that might exist down to an *ATM*-like profile of risk and mutational frequency, as well as all genes of a *PALB2*-like risk (OR = 5), down to a population MAF_{combined} of 0.0075% (1 in 6702) (Supplementary Table S13, available at <https://doi.org/10.1016/j.annonc.2022.09.152>).

Conclusions

This is the largest study to date of enriched *BRCA*-negative breast cancer probands, in which 7.4% were found to carry a BCSG pathogenic mutation. With appropriate correction for multiple tests, aside from ‘rediscovery’ of known BCSGs, neither the targeted DNA repair gene analysis nor the WES approach yielded robust additional BCSGs. Outstanding BCSGs of high penetrance (OR > 5) must be extremely rare. However, many genes up to a

CHEK2/ATM-like risk may exist. Whilst a WES study involving 5- to 10-fold more enriched breast cancer samples is feasible and would allow comprehensive survey of all such genes, there is no unanimous agreement regarding utility of clinical testing for genes that are so rare or of such modest penetrance.

ACKNOWLEDGEMENTS

We thank all the subjects and families that participated in the research. We thank those at the ICR, past and present, for their assistance in patient recruitment, sample management, and management of the sequencing facility. We are grateful to all the clinicians and counsellors in The Breast and Ovarian Cancer Susceptibility Collaboration who have contributed to the recruitment and collection of samples. The full list of contributors is provided in the Appendix.

FUNDING

This work was supported by Cancer Research UK [grants numbers C8620/A8372, C8620/A8857]; the Institute of Cancer Research (no grant number); NHS to the Institute of

Table 2. Population minor allele frequency at which genes with effect size equivalent to known breast cancer susceptibility genes are discoverable

Gene effect size	Population frequency detectable with 90% power			
	DNA repair gene analysis		WES analysis	
	Population MAF	Population frequency	Population MAF	Population frequency
<i>BRCA1</i> -like (OR = 10.6)	4.30E-05	1 in 11 628	5.59E-05	1 in 8945
<i>BRCA2</i> -like (OR = 5.9)	1.70E-04	1 in 2924	2.23E-04	1 in 2242
<i>PALB2</i> -like (OR = 5)	2.60E-04	1 in 1946	3.39E-04	1 in 1475
<i>CHEK2</i> -like (OR = 2.5)	1.90E-03	1 in 260	2.62E-03	1 in 191
<i>ATM</i> -like (OR = 2.1)	3.70E-03	1 in 135	4.98E-03	1 in 100
Gene effect size	Population frequency detectable with 80% power			
	DNA repair gene analysis		WES analysis	
	Population MAF	Population frequency	Population MAF	Population frequency
<i>BRCA1</i> -like (OR = 10.6)	3.56E-05	1 in 14 045	4.63E-05	1 in 10 799
<i>BRCA2</i> -like (OR = 5.9)	1.35E-04	1 in 3704	1.82E-04	1 in 2747
<i>PALB2</i> -like (OR = 5)	2.04E-04	1 in 2451	2.81E-04	1 in 1779
<i>CHEK2</i> -like (OR = 2.5)	1.54E-03	1 in 325	2.16E-03	1 in 231
<i>ATM</i> -like (OR = 2.1)	2.98E-03	1 in 168	4.18E-03	1 in 120

Gene population frequency (MAF_{combined}) for genes of equivalent effect size to *BRCA1*, *BRCA2*, *PALB2*, *CHEK2*, and *ATM* for which this study of 2135 enriched cases versus 51 377 controls had 90% and 80% power to detect an association. Bonferroni correction applied for undertaking targeted ($N = 286$ DNA repair genes) or exome-wide analysis ($N = 19\ 651$ genes). OR for *BRCA1*, *BRCA2*, *PALB2*, *CHEK2*, and *ATM* are derived from Dorling et al.²⁵ MAF, minor allele frequency; OR, odds ratio; WES, whole-exome sequencing.

Cancer Research and Royal Marsden as part of a joint entity referred to as the National Institute of Health Research Specialist Biomedical Research Centre for Cancer.

DISCLOSURE

The authors have declared no conflicts of interest.

REFERENCES

1. Cancer Research UK Cancer Incidence. Available at <https://www.cancerresearchuk.org/health-professional/cancer-statistics/statistics-by-cancer-type/breast-cancer#heading-Zero>.
2. Sung H, Ferlay J, Siegel RL, et al. Global cancer statistics 2020: GLOBOCAN estimates of incidence and mortality worldwide for 36 cancers in 185 countries. *CA Cancer J Clin*. 2021;71(3):209-249.
3. Lynch HT, Krush AJ, Lemon HM, Kaplan AR, Condit PT, Bottomley RH. Tumor variation in families with breast cancer. *JAMA*. 1972;222(13):1631-1635.
4. Collaborative Group on Hormonal Factors in Breast Cancer. Familial breast cancer: collaborative reanalysis of individual data from 52 epidemiological studies including 58,209 women with breast cancer and 101,986 women without the disease. *Lancet*. 2001;358(9291):1389-1399.
5. Turnbull C, Sud A, Houlston RS. Cancer genetics, precision prevention and a call to action. *Nat Genet*. 2018;50(9):1212-1218.
6. Lilyquist J, Ruddy KJ, Vachon CM, Couch FJ. Common genetic variation and breast cancer risk-past, present, and future. *Cancer Epidemiol Biomarkers Prev*. 2018;27(4):380-394.
7. Bahcall O. Common variation and heritability estimates for breast, ovarian and prostate cancers. *Nat Genet*. 2013. <https://doi.org/10.1038/ngicogs.1>.
8. Michailidou K, Lindstrom S, Dennis J, et al. Association analysis identifies 65 new breast cancer risk loci. *Nature*. 2017;551(7678):92-94.
9. Snape K, Ruark E, Tarpey P, et al. Predisposition gene identification in common cancers by exome sequencing: insights from familial breast cancer. *Breast Cancer Res Treat*. 2012;134(1):429-433.
10. Guo MH, Plummer L, Chan YM, Hirschhorn JN, Lippincott MF. Burden testing of rare variants identified through exome sequencing via publicly available control data. *Am J Hum Genet*. 2018;103(4):522-534.
11. Rahman N. Realizing the promise of cancer predisposition genes. *Nature*. 2014;505(7483):302-308.
12. Sondka Z, Bamford S, Cole CG, Ward SA, Dunham I, Forbes SA. The COSMIC cancer gene census: describing genetic dysfunction across all human cancers. *Nat Rev Cancer*. 2018;18(11):696-705.
13. Knijnenburg TA, Vockley JG, Chambwe N, et al. Genomic and molecular characterization of preterm birth. *Proc Natl Acad Sci U S A*. 2019;116(12):5819-5827.
14. Chakravarty D, Gao J, Phillips SM, et al. OncoKB: a precision oncology knowledge base. *JCO Precis Oncol*. 2017;2017:PO.17.00011.
15. Karczewski KJ, Francioli LC, Tiao G, et al. The mutational constraint spectrum quantified from variation in 141,456 humans. *Nature*. 2020;581(7809):434-443.
16. Richards S, Aziz N, Bale S, et al. Standards and guidelines for the interpretation of sequence variants: a joint consensus recommendation of the American College of Medical Genetics and Genomics and the Association for Molecular Pathology. *Genet Med*. 2015;17(5):405-424.
17. Cubuk C, Garrett A, Choi S, et al. Clinical likelihood ratios and balanced accuracy for 44 in silico tools against multiple large-scale functional assays of cancer susceptibility genes. *Genet Med*. 2021;23:2096-2104.
18. Ioannidis NM, Rothstein JH, Pejaver V, et al. REVEL: an ensemble method for predicting the pathogenicity of rare missense variants. *Am J Hum Genet*. 2016;99(4):877-885.
19. Wang Q, Dhindsa RS, Carss K, et al. Rare variant contribution to human disease in 281,104 UK Biobank exomes. *Nature*. 2021;597(7877):527-532.
20. Antoniou AC, Easton DF. Polygenic inheritance of breast cancer: implications for design of association studies. *Genet Epidemiol*. 2003;25(3):190-202.
21. Brewer HR, Jones ME, Schoemaker MJ, Ashworth A, Swerdlow AJ. Family history and risk of breast cancer: an analysis accounting for family structure. *Breast Cancer Res Treat*. 2017;165(1):193-200.
22. Kheirelseid EA, Jumustafa H, Miller N, et al. Bilateral breast cancer: analysis of incidence, outcome, survival and disease characteristics. *Breast Cancer Res Treat*. 2011;126(1):131-140.
23. Seal S, Thompson D, Renwick A, et al. Truncating mutations in the Fanconi anemia J gene BRIP1 are low-penetrance breast cancer susceptibility alleles. *Nat Genet*. 2006;38(11):1239-1241.
24. Roberts ME, Jackson SA, Susswein LR, et al. MSH6 and PMS2 germ-line pathogenic variants implicated in Lynch syndrome are associated with breast cancer. *Genet Med*. 2018;20(10):1167-1174.
25. Dorling L, Carvalho S, Allen J, et al. Breast cancer risk genes - association analysis in more than 113,000 women. *N Engl J Med*. 2021;384(5):428-439.
26. Ruark E, Snape K, Humburg P, et al. Mosaic PPM1D mutations are associated with predisposition to breast and ovarian cancer. *Nature*. 2013;493(7432):406-410.
27. Kim B, Won D, Lee ST, Choi JR. Somatic mosaic truncating mutations of PPM1D in blood can result from expansion of a mutant clone under selective pressure of chemotherapy. *PLoS One*. 2019;14(6):e0217521.
28. Pharoah PDP, Song H, Dicks E, et al. PPM1D mosaic truncating variants in ovarian cancer cases may be treatment-related somatic mutations. *J Natl Cancer Inst*. 2016;108(3):djv347.
29. Ashburner M, Ball CA, Blake JA, et al. Gene ontology: tool for the unification of biology. The Gene Ontology Consortium. *Nat Genet*. 2000;25(1):25-29.
30. Gene Ontology Consortium. The gene ontology resource: enriching a gold mine. *Nucleic Acids Res*. 2021;49(D1):D325-D334.
31. Hall JM, Lee MK, Newman B, et al. Linkage of early-onset familial breast cancer to chromosome 17q21. *Science*. 1990;250(4988):1684-1689.
32. Wooster R, Neuhausen SL, Mangion J, et al. Localization of a breast cancer susceptibility gene, BRCA2, to chromosome 13q12-13. *Science*. 1994;265(5181):2088-2090.
33. Wooster R, Bignell G, Lancaster J, et al. Identification of the breast cancer susceptibility gene BRCA2. *Nature*. 1995;378(6559):789-792.
34. Miki Y, Swensen J, Shattuck-Eidens D, et al. A strong candidate for the breast and ovarian cancer susceptibility gene BRCA1. *Science*. 1994;266(5182):66-71.
35. Smith P, McGuffog L, Easton DF, et al. A genome wide linkage search for breast cancer susceptibility genes. *Genes Chromosomes Cancer*. 2006;45(7):646-655.
36. Meijers-Heijboer H, van den Ouweland A, Klijn J, et al. Low-penetrance susceptibility to breast cancer due to CHEK2(*)1100delC in noncarriers of BRCA1 or BRCA2 mutations. *Nat Genet*. 2002;31(1):55-59.
37. Thompson D, Duedal S, Kirner J, et al. Cancer risks and mortality in heterozygous ATM mutation carriers. *J Natl Cancer Inst*. 2005;97(11):813-822.
38. Rahman N, Seal S, Thompson D, et al. PALB2, which encodes a BRCA2-interacting protein, is a breast cancer susceptibility gene. *Nat Genet*. 2007;39(2):165-167.
39. Lu HM, Li S, Black MH, et al. Association of breast and ovarian cancers with predisposition genes identified by large-scale sequencing. *JAMA Oncol*. 2019;5(1):51-57.
40. Huang KL, Mashl RJ, Wu Y, et al. Pathogenic germline variants in 10,389 adult cancers. *Cell*. 2018;173(2):355-370 e14.
41. Lu C, Xie M, Wendl MC, et al. Patterns and functional implications of rare germline variants across 12 cancer types. *Nat Commun*. 2015;6:10086.

APPENDIX

Breast and Ovarian Cancer Susceptibility Collaboration

The following individuals are part of Breast and Ovarian Cancer Susceptibility Collaboration:

- A. Ardern-Jones¹, J. Adlard², M. Ahmed³, G. Attard⁴, K. Bailey⁵, E. Bancroft¹, C. Bardsley⁶, D. Barton⁷, M. Bartlett⁸, J. Barwell⁹, L. Baxter⁶, R. Belk¹⁰, J. Berg¹¹, B. Bernhard⁸, T. Bishop², L. Boyes¹², N. Bradshaw¹³, A.F. Brady⁸, S. Brant¹⁴, C. Brewer¹⁴, G. Brice⁴, G. Bromilow¹⁴, C. Brooks¹⁵, A. Bruce¹⁶, B. Bulman¹⁰, L. Burgess¹², J. Campbell¹⁷, N. Canham⁸, B. Castle³, R. Cetnarskyj¹⁷, C. Chapman¹², O. Claber⁸, N. Coates¹⁹, T. Cole¹², A. Collins³, J. Cook⁶, S. Coulson¹⁵, G. Crawford³, D. Cruger¹², C. Cummings⁸, L. D'Mello¹, R. Davidson¹³, L. Day³, B. Dell³, C. Dolling¹², A. Donaldson¹⁹, H. Dorkins⁸, F. Douglas¹⁸, S. Downing²⁰, S. Drummond¹⁷, C. Dubras³, J. Dunlop¹¹, S. Durrell²¹, D. Eccles³, C. Eddy⁴, M. Edwards¹⁷, E. Edwards³, J. Edwardson²², R. Eeles¹, I. Ellis¹⁵, F. Elmslie⁴, G. Evans¹⁰, B. Gibbons²², C. Gardiner³, N. Ghali⁸, C. Giblin¹⁰, S. Gibson¹³, S. Goff⁴, S. Goodman¹⁴, D. Goudie¹¹, L. Greenhalgh¹⁵, J. Grier⁵, H. Gregory²², S. Halliday¹, R. Hardy⁷, C. Hartigan⁹, T. Heaton¹², A. Henderson¹⁸, C. Higgins¹⁵, S. Hodgson⁴, T. Homfray⁴, D. Horrigan⁵, C. Houghton¹⁵, R.S. Houlston¹, L. Hughes²⁰, V. Hunt¹⁴, L. Irvine¹³, L. Izatt²³, C. Jacobs²³, S. James²², M. James¹⁹, L. Jeffers⁷, I. Jobson¹⁸, W. Jones⁴, M.J. Kennedy⁵, S. Kenwick²⁰, C. Kightley¹⁵, C. Kirk⁷, E. Kirk⁵, E. Kivuva¹⁴, K. Kohut¹, M. Kosicka-Slawinska⁸, A. Kulkarni⁸, A. Kumar¹⁶, F. Lalloo¹⁰, N. Lambord¹⁴, C. Langman²³, P. Leonard¹², S. Levene²³, S. Locker⁵, P. Logan⁷, M. Longmuir¹³, A. Lucassen³, V. Lyus⁸, A. Magee⁷, A. Male¹⁶, S. Mansour⁴, D. McBride³, E. McCann⁵, V. McConnell⁷, M. McEntagart⁴, C. McKeown¹², L. McLeish¹¹, D. McLeod²¹, A. Melville⁸, L. Mercer⁶, C. Mercer³, Z. Miedzybrodzka²², A. Mitra¹, P. J. Morrison⁷, V. Murday⁴, A. Murray⁵, K. Myhill¹, J. Myring¹¹, E. O'Hara², J. Paterson²⁰, P. Pearson¹⁷, G. Pichert²³, K. Platt⁵, M. Porteous¹⁷, C. Pottinger¹⁵, S. Price²¹, L. Protheroe⁵, S. Pugh¹⁹, O. Quarrell⁶, K. Randhawa⁸, C. Riddick⁵, L. Robertson¹, A. Robinson²⁰, V. Roffey-Johnson⁹, M. Rogers⁵, S. Rose⁸, S. Rowe¹⁰, A. Schofield²², N. Rahman¹, S. Saya¹, G. Scott¹³, J. Scott⁷, A. Searle¹⁴, S. Shanley¹, S. Sharif¹², A. Shaw²³, J. Shaw¹⁵, J. Shea-Simonds³, L. Side¹⁶, J. Sillibourne³, K. Simon¹⁶, S. Simpson²², S. Slater¹⁷, S. Smalley³, K. Smith¹⁹, L. Snadden¹³, K. Snape¹, J. Soloway¹⁶, Y. Stait¹⁴, B. Stayner²¹, M. Steel¹⁷, C. Steel¹¹, H. Stewart²¹, D. Stirling¹⁷, M. Thomas¹⁴, S. Thomas¹, S. Tomkins¹⁹, H. Turner⁵, A. Vandersteens⁸, E. Wakeling⁸, F. Waldrup²⁰, L. Walker²¹, C. Watt¹³, S. Watts²³, A. Webber¹⁵, C. Whyte¹⁷, J. Wiggins¹, E. Williams¹⁶, L. Winchester⁴
1. Department of Cancer Genetics, Royal Marsden NHS Foundation Trust, Sutton, UK.
 2. Yorkshire Regional Genetics Service, Chapel Allerton Hospital, Leeds, UK.
 3. Faculty of Medicine, University of Southampton, Southampton University Hospitals NHS Trust, UK.
 4. SW Thames Regional Genetics Service, St George's Hospital, London, UK.
 5. All Wales Medical Genetics Service, University Hospital of Wales, Cardiff, UK.
 6. Sheffield Regional Genetics Service, Sheffield Children's NHS Foundation Trust, UK.
 7. Northern Ireland Regional Genetics Service, Belfast HSC Trust, Department of Medical Genetics, Queen's University Belfast, UK.
 8. NW Thames Regional Genetics Service, Kennedy Galton Centre, London, UK.
 9. Leicestershire Genetics Centre, University Hospitals of Leicester NHS Trust, UK.
 10. Genetic Medicine, Manchester Academic Health Science Centre, St. Mary's Hospital, Manchester, UK.
 11. Human Genetics, Division of Medical Sciences, University of Dundee, UK
 12. West Midlands Regional Genetics Service, Birmingham Women's Hospital, Birmingham, UK.
 13. West of Scotland Regional Genetics Service, Laboratory Medicine, Southern General Hospital, Glasgow, UK.
 14. Peninsula Regional Genetics Service, Royal Devon & Exeter Hospital, Exeter, UK.
 15. Merseyside and Cheshire Clinical Genetics Service, Liverpool Women's NHS Foundation Trust, Liverpool, UK.
 16. NE Thames Regional Genetics Service, Great Ormond St Hospital, London, UK.
 17. South East of Scotland Clinical Genetics Service, Western General Hospital, Edinburgh, UK.
 18. Northern Genetics Service, Newcastle upon Tyne Hospitals NHS Foundation Trust, UK.
 19. South Western Regional Genetics Service, University Hospitals of Bristol NHS Foundation Trust, UK.
 20. East Anglian Regional Genetics Service, Cambridge University Hospitals NHS Foundation Trust, UK.
 21. Oxford Regional Genetics Service, Oxford University Hospitals NHS Trust, Oxford, UK.
 22. University of Aberdeen and North of Scotland Clinical Genetics Service, Aberdeen Royal Infirmary, UK.
 23. SE Thames Regional Genetics Service, Guy's and St Thomas NHS Foundation Trust, London, UK.