ESMO GOOD SCIENCE BETTER MEDICINE BEST PRACTICE

ANNALS OF ONCOLOGY driving innovation in oncology

ORIGINAL ARTICLE

Population-based germline breast cancer gene association studies and meta-analysis to inform wider mainstream testing

C. F. Rowlands[1], S. Allen[1], J. Balmaña[2,3], S. M. Domchek[4], D. G. Evans[5], H. Hanson[6,7], N. Hoogerbrugge[8], P. A. James[9], K. L. Nathanson[10,11], M. Robson[12,13], M. Tischkowitz[14], W. D. Foulkes[15,16,17] & C. Turnbull[1]*

[1]Division of Genetics and Epidemiology, The Institute of Cancer Research, London, UK; [2]Hereditary Cancer Genetics Group, Vall d'Hebron Institute of Oncology (VHIO), Barcelona; [3]Medical Oncology Department, Hospital Universitari Vall d'Hebron, Vall d'Hebron Hospital Campus, Barcelona, Spain; [4]Basser Center for BRCA, Abramson Cancer Center, University of Pennsylvania, Philadelphia, USA; [5]Manchester Academic Health Science Centre, Division of Evolution and Genomic Sciences, School of Biological Sciences, Faculty of Biology, Medicine and Health, The University of Manchester, Manchester; [6]Department of Clinical and Biomedical Sciences, University of Exeter Medical School, Exeter; [7]Peninsula Regional Genetics Service, Royal Devon University Healthcare NHS Foundation Trust, Exeter, UK; [8]Department of Human Genetics, Radboud University Medical Center, Nijmegen, The Netherlands; [9]Parkville Familial Cancer Centre, The Royal Melbourne Hospital and Peter MacCallum Cancer Centre, Melbourne, Australia; [10]Division of Translational Medicine and Human Genetics, Department of Medicine, Perelman School of Medicine, University of Pennsylvania, Philadelphia; [11]Abramson Cancer Center, Perelman School of Medicine, Philadelphia; [12]Department of Medicine, Memorial Sloan Kettering Cancer Center, New York; [13]Weill Cornell Medical College, New York, USA; [14]Department of Medical Genetics, National Institute for Health Research Cambridge Biomedical Research Centre, University of Cambridge, Cambridge, UK; Departments of [15]Human Genetics, [16]Oncology and [17]Medicine, McGill University, Montréal, Canada

Available online XXX

Background: Germline genetic testing, previously restricted to familial and young-onset breast cancer, is now offered increasingly broadly to patients with 'population-type' breast cancer in mainstream oncology clinics, with wide variation in the genes included.

Patients and methods: Weighted meta-analysis was carried out for three population-based case—control studies (BRIDGES, CARRIERS and UK Biobank) comprising in total 101 397 women with breast cancer and 312 944 women without breast cancer, to quantify 37 putative breast cancer susceptibility genes (BCSGs) for the frequency of pathogenic variants (PVs) in unselected, 'population-type' breast cancer cases and their association with breast cancer and its subtypes.

Results: Meta-analysed odds ratios (ORs) and frequencies of PVs in 'population-type' breast cancer cases were generated for *BRCA1* (OR 8.73, 95% confidence interval (CI) 7.47-10.20; 1 in 101), *BRCA2* (OR 5.68, 95% CI 5.13-6.30; 1 in 68) and *PALB2* (OR 4.30, 95% CI 3.68-5.03; 1 in 187). For both *CHEK2* (OR 2.40, 95% CI 2.21-2.62; 1 in 73) and *ATM* (OR 2.16, 95% CI 1.93-2.41; 1 in 132) subgroup analysis showed a stronger association with oestrogen receptor-positive disease. The magnitude of association and frequency of PVs were low for *RAD51C* (OR 1.53, 95% CI 1.29-2.04; 1 in 913), *RAD51D* (OR 1.76, 95% CI 1.29-2.41; 1 in 1079) and *BARD1* (OR 2.34, 95% CI 1.85-2.97; 1 in 672); frequencies and associations were higher when the analysis was restricted to triple-negative breast cancers. The PV frequency in 'population-type' breast cancer cases was very low for 'syndromic' BCSGs *TP53* (1 in 1844), *STK11* (1 in 11 525), *CDH1* (1 in 2668), *PTEN* (1 in 3755) and *NF1* (1 in 1470), with metrics of association also modest ranging from OR 3.62 (95% CI 1.98-6.61) for *TP53* down to OR 1.60 (95% CI 0.48-5.30) for *STK11*.

Conclusions: These metrics reflecting 'population-type' breast cancer will be informative in defining the appropriate gene set as we continue to expand to germline testing to an increasingly unselected group of breast cancer cases.

Key words: meta-analysis, breast cancer, case—control, mainstream genetic testing, multigene panel testing

INTRODUCTION

Linkage analysis examining multicase families led to the identification of 'first-wave' breast cancer susceptibility genes (BCSGs) *BRCA1* in 1994 and *BRCA2* shortly thereafter.[1,2] Failure of linkage analysis in other unexplained large breast cancer pedigrees suggested that no additional BCSGs existed of equivalent prevalence—risk profile to *BRCA*1/2. The methodology moved to the large-scale mutational case—control screening experiments of candidate DNA repair genes, leading to the identification in the 2000s of *PALB2*, *ATM*, *CHEK2*, *BARD1*, *RAD51C* and *RAD51D* as BCSGs of weaker prevalence—risk profile.[3-8] In parallel, breast cancer was implicated as part of various 'syndromes' of pleomorphic

*Correspondence to: Prof. Clare Turnbull, Division of Genetics and Epidemiology, The Institute of Cancer Research, Cotswold Rd, Sutton, London, SM2 5NG, UK. Tel: +44208 722 4485

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

susceptibility to very rare cancers, linked to genes such as *TP53* and *CDH1*, as well as *PTEN*, *NF1* and *STK11* in which distinctive noncancer phenotypes are also described.

Until recently, estimates of risk for pathogenic variants (PVs) in these genes were derived from families ascertained based on multiple cases of closely-related, unusual or young-onset cancers, with the accordant concern of upward bias in these estimates. There was, therefore, a welcome reception in 2021 for two large collaborative analyses (BRIDGES and CARRIERS), each amalgamating multiple population-based series of breast cancer cases and paired controls, and analysing the frequency of PVs in multiple (putative) BCSGs.[9,10] However, these studies differed in their technical and analytical approaches. In particular, BRIDGES only reported on protein-truncating variants (PTVs), which are a subset of all PVs. These methodological distinctions, along with the overall rarity of PVs, meant that for some BCSGs there were clinically important differences between the two studies regarding reported association with breast cancer, magnitudes of association (effect sizes) and PV frequency. Also now widely published are analyses from population-based cohort studies such as UK Biobank (UKB), which includes >20 000 cases of female breast cancer.

Herein we present a comparison across BRIDGES, CARRIERS and UKB of the frequencies in women with breast cancer from population-based studies (cases) and in women without breast cancer (controls) of PVs in 37 (putative) BCSGs, harmonising the datasets to incorporate PTVs with non-PTV PVs (as this combined total is the metric most relevant to clinical testing). We then undertake a weighted meta-analysis of the three datasets, involving a total of 101 397 breast cancer cases and 312 944 controls, to generate for each gene weighted-average estimates for breast cancer association [odds ratios (ORs)] and PV frequencies for cases and (population) controls.

METHODS

Studies and patients

We included summary data on 33 breast cancer-associated genes from 48 826 female patients with breast cancer (including invasive and *in situ* disease) and 50 703 controls from the BRIDGES consortium from 30 contributing BCAC (Breast Cancer Association Consortium) studies recruiting patients unselected for family history. We included summary data for 28 breast cancer-associated genes from 32 247 female patients with breast cancer (including invasive and *in situ* disease) and 32 544 controls from the CARRIERS Consortium from 12 contributing studies recruiting patients unselected for family history. We utilised individual-level data on these 37 breast cancer-associated genes included in the aforementioned studies from female participants from UKB, assigning as breast cancer cases those with any past assignment of invasive or *in situ* breast cancer from cancer registrations or self-report and assigning those with no history of breast disease as controls, totalling 20 324 cases of female breast cancer and 229 697 controls[11] (see Supplementary Methods, available at https://doi.org/10.1016/j.annonc.2024.07.244). The three datasets utilised distinct targeted DNA capture kits, which are detailed further in the Supplementary Methods, available at https://doi.org/10.1016/j.annonc.2024.07.244.

Bioinformatic analysis

For UKB, we extracted variants for the 37 genes of interest from the population variant call format file, retaining exonic variants extending ±25 bp into introns for which there was coverage of >10 reads at the variant position in at least 90% of all UKB samples, to ensure that only variants at positions with high callability were considered. We utilised the consensus or longest transcript used in BRIDGES/CARRIERS. As per thresholds applied in CARRIERS, variants of low variant allele frequency (VAF) were excluded due to the potential for clonal haematopoietic origin (VAF <0.3/VAF >0.7 for *TP53* and *NF1*, VAF <0.05/VAF >0.95 for other genes). We ascribed variants as pathogenic based on Ensembl Variant Effect Predictor annotation predicting protein truncation (PTV), discounting those in the final exon or penultimate donor splice region (with some exceptions, detailed in Supplementary Methods, available at https://doi.org/10.1016/j.annonc.2024.07.244). We included additional non-PTV PVs based on ClinVar annotation of pathogenic/likely pathogenic (≥2 stars) provided the variant had no conflicting interpretations. For BRIDGES, we extracted from the publicly available variant-level summary data all pathogenic missense variants based on ClinVar annotation of pathogenic/likely pathogenic (≥2 stars). We did not include lower-risk variants assigned as 'risk factor' or of 'conflicting interpretations of pathogenicity', such as *CHEK2* I157T or S428F. Additional details are presented in Supplementary Methods, available at https://doi.org/10.1016/j.annonc.2024.07.244.

Statistical methods

For UKB, we estimated ORs for breast cancer by logistic regression, adjusting for ethnicity and age considering (i) all PVs and (ii) just PTVs. We calculated a gene-specific adjustment from the BRIDGES PTV data by comparing directly estimated OR and multiple logistic regression-adjusted OR (Supplementary Figure S1, available at https://doi.org/10.1016/j.annonc.2024.07.244). Although of small magnitude, this adjustment was applied to the BRIDGES missense PV data. We then integrated summary statistics for the three constituent studies to generate combined weighted average ORs for each gene, using a fixed effects inverse variance approach. For all PVs (PTVs and non-PTVs) this comprised BRIDGES, CARRIERS and UKB (101 397 cases and 312 944 controls). For the analysis of PTVs, this included only BRIDGES and UKB. Lower and upper 90% and 95% confidence intervals (CIs) were calculated from the standard error of the OR estimate.

RESULTS

Supplementary Table S1, available at https://doi.org/10.1016/j.annonc.2024.07.244 illustrates the characteristics of the three contributing studies, from which a total of 101 397 breast

Table 1. Pathogenic variant frequencies for 13 genes in cases and controls for constituent studies and the combined meta-analysis

Gene		Cases					Controls				
		BRIDGES	CARRIERS	UKB	p_{het}	Combined	BRIDGES	CARRIERS	UKB	p_{het}	Combined
BCSGs of higher penetrance	***BRCA1***	612 (1.25%)	275 (0.85%)	112 (0.55%)	2.49×10^{-18}	**999 (0.99%; 101)**	77 (0.15%)	37 (0.11%)	162 (0.070%)	4.43×10^{-8}	**276 (0.09%; 1134)**
	BRCA2	808 (1.65%)	417 (1.29%)	274 (1.35%)	3.71×10^{-5}	**1499 (1.48%; 68)**	150 (0.30%)	78 (0.24%)	555 (0.24%)	0.08	**783 (0.25%; 400)**
	PALB2	274 (0.56%)	148 (0.46%)	121 (0.60%)	0.06	**543 (0.54%; 187)**	55 (0.10%)	38 (0.12%)	338 (0.15%)	0.06	**431 (0.14%; 726)**
BCSGs of more moderate penetrance	***CHEK2***	761 (1.56%)	349 (1.08%)	275 (1.35%)	7.79×10^{-8}	**1385 (1.37%; 73)**	346 (0.68%)	138 (0.42%)	1338 (0.58%)	1.09×10^{-5}	**1822 (0.58%; 172)**
	ATM	350 (0.72%)	253 (0.78%)	164 (0.80%)	0.36	**767 (0.76%; 132)**	175 (0.35%)	134 (0.41%)	781 (0.34%)	0.12	**1090 (0.35%; 287)**
	BARD1	62 (0.13%)	49 (0.15%)	40 (0.20%)	0.09	**151 (0.15%; 672)**	32 (0.063%)	35 (0.10%)	126 (0.055%)	1.61×10^{-3}	**193 (0.06%; 1621)**
	RAD51C	56 (0.11%)	41 (0.13%)	14 (0.069%)	0.13	**111 (0.11%; 913)**	26 (0.051%)	35 (0.10%)	111 (0.048%)	1.04×10^{-4}	**172 (0.05%; 1819)**
	RAD51D	51 (0.10%)	26 (0.080%)	17 (0.084%)	0.49	**94 (0.09%; 1079)**	25 (0.049%)	14 (0.043%)	110 (0.048%)	0.91	**149 (0.05%; 2100)**
'Syndromic' BCSGs	***TP53***	34 (0.070%)	19 (0.059%)	2 (0.0098%)	8.01×10^{-3}	**55 (0.05%; 1844)**	9 (0.018%)	2 (0.0061%)	12 (0.0052%)	0.01	**23 (0.01%; 13 606)**
	PTEN	16 (0.033%)	8 (0.025%)	3 (0.015%)	0.41	**27 (0.03%; 3755)**	7 (0.014%)	3 (0.0092%)	11 (0.0048%)	0.07	**21 (0.01%; 14 902)**
	NF1	37 (0.076%)	19 (0.059%)	13 (0.064%)	0.65	**69 (0.07%; 1470)**	21 (0.041%)	11 (0.034%)	70 (0.030%)	0.46	**102 (0.03%; 3068)**
	STK11	6 (0.012%)	NA	0 (0%)	0.19	**6 (0.01%; 11525)**	5 (0.0099%)	NA	1 (0.00044%)	9.85×10^{-4}	**6 (0%; 46 733)**
	CDH1	12 (0.025%)	17 (0.053%)	9 (0.044%)	0.11	**38 (0.04%; 2668)**	14 (0.028%)	6 (0.018%)	27 (0.012%)	0.03	**47 (0.02%; 6658)**

Carrier counts are given for each cohort, dataset and gene, with the respective percentage of total samples given in brackets. Data are shown for the combined variant cohort comprising PTVs + pathogenic missense (for BRIDGES) or non-PTV-PVs (CARRIERS and UKB). The combined counts detail both the percentage of total samples carrying PVs followed by a "1 in..." figure, rounded to the nearest whole number of individuals, in brackets. p_{het}-values were derived from a Fisher exact test where there were only two contributing datasets and at least one of the carrier counts was <6; otherwise, p-values were derived from a chi-squared test (Bonferroni-corrected p-value threshold of significance $<2.29 \times 10^{-4}$) Combined percentages are calculated against the sum of the total cases or controls in those constituent studies that included the respective gene. The BRIDGES dataset comprises 48 826 cases and 50 704 controls; CARRIERS comprises 32 247 cases and 32 544 controls; UKB comprises 20 324 cases and 229 697 controls. The Combined total comprises 101 397 cases and 312 944 controls.

BCSG, breast cancer susceptibility gene; NA, not applicable; PTV, protein-truncating variant; PV, pathogenic variant; UKB, UK Biobank.

cancer cases and 312 944 controls were included. In Table 1, we present the PV frequencies for breast cancer cases (BC PV frequency) and controls (control PV frequency) for each of the three contributing studies, as well as their combined weighted averages. This table includes 13 genes widely recognized as BCSGs, along with *P* values for heterogeneity. In Table 2, we present the ORs for PTVs alone and all PVs for each contributing study and for the combined weighted meta-analysis. In Figure 1, we illustrate these ORs on a logarithmic scale, along with their corresponding CIs. Additional data on ORs and PV frequencies for all 37 genes included in BRIDGES and/or CARRIERS and detailed frequency data are presented in Supplementary Tables S2A, B, S3 and S4, available at https://doi.org/10.1016/j.annonc.2024.07.244.

Breast cancer susceptibility genes of higher penetrance (risk)

For *BRCA1*, there was no evidence of significant heterogeneity in the OR estimates across the three contributing studies, with the weighted average being higher for PTVs (combined OR 9.85, 95% CI 8.13-11.94) compared with all PVs (combined OR 8.73, 95% CI 7.47-10.20). However, there was a marked variation in PV frequency between the three contributing studies, with control PV frequencies being 0.15% in BRIDGES, 0.11% in CARRIERS and 0.07% in UKB ($P_{het} = 4.43 \times 10^{-8}$; Table 1). This likely represents technical differences between the studies, thus being consistent and nondifferential across cases and controls. Weighted average frequencies for all PVs were estimated to be 0.99% (1 in 101) in breast cancer cases and 0.09% for controls (1 in 1134).

For *BRCA2*, the frequency of PVs in controls was consistent across the three studies, with little difference between OR estimates for PTVs alone and all PVs. The combined weighted average across the studies for all PVs was OR 5.68 (95% CI 5.13-6.30), with a BC PV frequency of 1.48% (1 in 68) and a control PV frequency of 0.25% (1 in 400).

For *PALB2*, the published OR for CARRIERS was 3.83 (95% CI 2.68-5.63), while for BRIDGES this was OR 5.02 (95% CI 3.73-6.76). However, there was no evidence of significant heterogeneity between the studies in terms of variant frequencies, with the estimate of association for UKB being intermediary between BRIDGES and CARRIERS (OR 4.14, 95% CI 3.35-5.10). The combined weighted average for all PVs was estimated to be OR 4.30 (95% CI 3.68-5.03), with average PV frequencies of 0.54% (1 in 187) and 0.14% (1 in 726) among studies in breast cancer cases and controls, respectively.

Thus the frequency of all PVs for *BRCA1*, *BRCA2* and *PALB2* combined is 3.00% (1 in 33) in breast cancer cases and 0.48% (1 in 210) in controls (i.e. in the general population).

Breast cancer susceptibility genes of more moderate penetrance (risk)

There was some variation between studies for PV frequencies in *CHEK2*. The frequencies for all PVs in breast cancer cases were 1.56% (BRIDGES), 1.08% (CARRIERS), and

Table 2. ORs for 13 genes for constituent studies and from a combined meta-analysis

Genes	PTVs			PTVs + non-PTV-PVs					
	Odds ratios			Odds ratios				Heterogeneity metrics	
	BRIDGES	UKB	Combined	BRIDGES	CARRIERS	UKB	Combined	Cochran's *Q*	I^2
BRCA1	10.57 (8.02-13.93)	9.22 (7.04-12.06)	9.85 (8.13-11.94)	9.24 (7.26-11.76)	7.62 (5.33-11.27)	8.72 (6.83-11.13)	8.73 (7.47-10.20)	0.72	0
BRCA2	5.85 (4.85-7.06)	5.79 (4.98-6.75)	5.82 (5.17-6.55)	5.60 (4.68-6.69)	5.23 (4.09-6.77)	5.91 (5.10-6.84)	5.68 (5.13-6.30)	0.71	0
PALB2	5.02 (3.73-6.76)	4.05 (3.25-5.05)	4.37 (3.66-5.21)	5.02 (3.73-6.76)	3.83 (2.68-5.63)	4.14 (3.35-5.10)	4.30 (3.68-5.03)	1.55	0
CHEK2	2.54 (2.21-2.91)	2.35 (2.04-2.69)	2.44 (2.21-2.69)	2.46 (2.15-2.80)	2.47 (2.02-3.05)	2.33 (2.04-2.65)	2.40 (2.21-2.62)	0.41	0
ATM	2.10 (1.71-2.57)	2.24 (1.83-2.73)	2.17 (1.88-2.50)	2.14 (1.78-2.58)	1.82 (1.46-2.27)	2.40 (2.03-2.85)	2.16 (1.93-2.41)	3.81	47.55
ARD1	2.09 (1.35-3.23)	3.50 (2.43-5.03)	2.83 (2.14-3.74)	2.09 (1.35-3.23)	1.37 (0.87-2.16)	3.54 (2.47-5.07)	2.34 (1.85-2.97)	10.71	81.33
RAD51C	1.93 (1.2-3.11)	1.19 (0.60-2.36)	1.65 (1.11-2.44)	1.97 (1.23-3.16)	1.20 (0.75-1.93)	1.51 (0.86-2.64)	1.53 (1.15-2.04)	2.13	6.14
RAD51D	1.80 (1.11-2.93)	1.74 (1.04-2.91)	1.77 (1.24-2.52)	1.80 (1.11-2.93)	1.72 (0.88-3.51)	1.74 (1.04-2.91)	1.76 (1.29-2.41)	0.01	0
TP53	3.06 (0.63-14.91)	1.59 (0.19-12.99)	2.41 (0.68-8.54)	3.30 (1.58-6.89)	9.59 (2.23-41.19)	1.88 (0.42-8.45)	3.62 (1.98-6.61)	2.50	20.15
PTEN	2.25 (0.85-6.00)	3.51 (0.39-31.71)	2.42 (0.99-5.91)	2.20 (0.89-5.44)	2.69 (0.71-10.15)	3.69 (1.02-13.34)	2.63 (1.38-5.02)	0.41	0
NF1	1.76 (0.96-3.21)	2.15 (1.16-4.00)	1.94 (1.26-2.99)	1.70 (0.98-2.93)	1.93 (0.91-4.31)	2.41 (1.42-4.07)	2.01 (1.43-2.83)	0.83	0
STK11	1.60 (0.48-5.28)	0.0025 (9.77x 10^{-41} - 6.58x10^{34})	1.60 (0.48-5.30)	1.60 (0.48-5.28)	NA	0.0025 (9.77x 10^{-41}-6.58x10^{34})	1.60* (0.48-5.30)	NA	NA
CDH1	0.86 (0.37-1.98)	3.83 (1.63-9.00)	1.79 (0.98-3.26)	0.81 (0.36-1.78)	2.50 (1.01-7.07)	4.11 (1.92-8.81)	2.01 (1.25-3.24)	8.70	77.02

Odds ratios for 13 genes for constituent studies and following combined meta-analysis. Odds ratios have been provided with the lower and upper 95% confidence intervals. Heterogeneity estimates are presented where genes were analysed in all three studies and comprise Cochran's Q (a chi-squared distributed measure of heterogeneity) and I^2 (a measure of heterogeneity more robust than Cochran's Q at low n, evaluating the percentage of variability in OR estimates due to heterogeneity rather than sampling error. I^2 may be broadly interpreted as follows: 0-40% = may be unimportant; 30-60% = may represent moderate heterogeneity; 50-90% = may represent substantial heterogeneity; 75-100% = considerable heterogeneity). The combined analyses for PTVs include BRIDGES and UKB only.

NA, not applicable; OR, odds ratio; PTV, protein-truncating variant; PV, pathogenic variant; UKB, UK Biobank.

*The combined meta-analysis for *STK11* uses data from BRIDGES and UKB only.

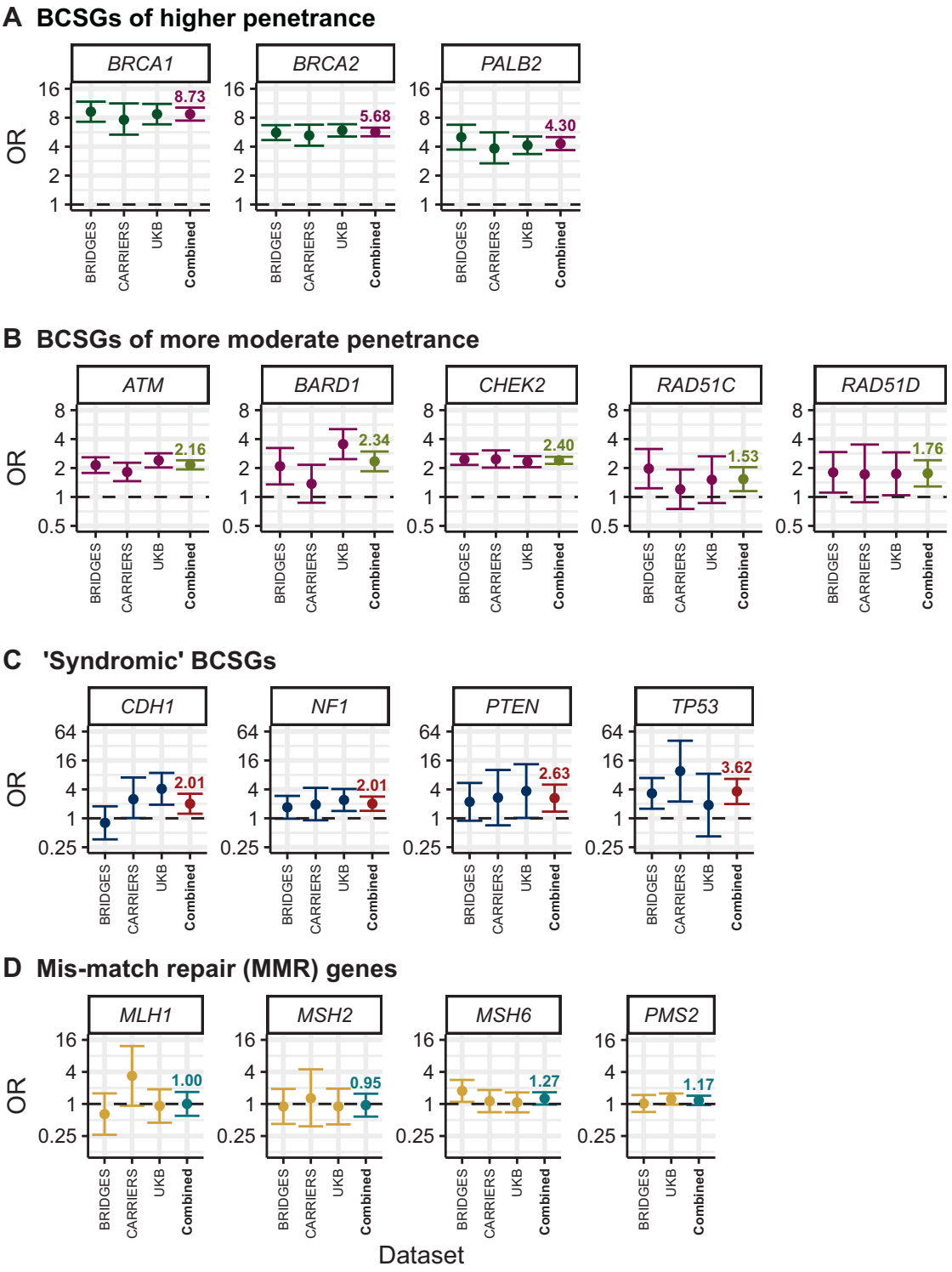

Figure 1. Logarithmic-scale OR towards breast cancer for pathogenic variants in selected genes, compared across each study and for the combined meta-analysis. (A) BCSGs of higher penetrance (*BRCA1, BRCA2,* and *PALB2*). (B) BCSGs of more moderate penetrance (*ATM, BARD1, CHEK2, RAD51C,* and *RAD51D*). (C) 'Syndromic' BCSGs (*CDH1, NF1, PTEN,* and *TP53*). (D) Mis-match repair genes (*MLH1, MSH2, MSH6* and *PMS2*). A reference boundary set to OR 1 is displayed on all plots (dashed line). BCSG, breast cancer susceptibility gene, OR, odds ratio. 95% confidence intervals are displayed for each OR.

1.35% (UKB), whereas in controls, these were 0.68%, 0.42%, and 0.58%, respectively ($P_{het} = 1.09 \times 10^{-5}$; Table 2). This resulted in weighted average PV frequency estimates for all *CHEK2* PVs of 1.37% in breast cancer cases and 0.58% in controls. However, these frequency differences across studies were nondifferential, that is, ORs for *CHEK2* were highly consistent across the three studies and very similar for PTVs and all PVs: weighted average for all PVs was OR 2.40 (95% CI 2.21-2.62) and for just the 1100delC variant was OR 2.54 (95% CI 2.21-2.62; BRIDGES and UKB only).

https://doi.org/10.1016/j.annonc.2024.07.244

For *ATM*, there was some heterogeneity in ORs, with the published CARRIERS association estimate being OR 1.82 (95% CI 1.46-2.27), whereas the association metrics were higher for BRIDGES and UKB, yielding a weighted average for all PVs of OR 2.16 (95% CI 1.93-2.41) and PV frequency of 0.76% and 0.35% in breast cancer cases and controls. Thus PVs in *CHEK2* and *ATM* together are present in 2.12% (1 in 47) of breast cancer cases and at a frequency of 0.93% (1 in 108) in controls (i.e. in the general population).

Breast cancer susceptibility genes more strongly associated with triple-negative breast cancer

For *RAD51D*, the PV frequencies and association metrics were similar in the three studies with a weighted average of OR 1.76 (95% CI 1.29-2.41). For both *BARD1* and *RAD51C*, there was some heterogeneity in the control PV frequency, which translated across into variation in association metrics. For *RAD51C*, the ORs for all PVs were as follows: CARRIERS OR 1.20 (95% CI 0.75-1.93), UKB OR 1.51 (95% CI 0.86-2.64) and BRIDGES OR 1.97 (95% CI 1.23-3.16). This resulted in a weighted average OR of 1.53 (95% CI 1.15-2.04). For *BARD1,* the weighted average OR for all PVs was 2.34 (95% CI 1.85-2.97), with the highest OR in UKB (OR 3.54, 95% CI 2.47-5.07) and the lowest in CARRIERS (OR 1.37, 95% CI 0.87-2.16).

From the combined analysis of BRIDGES and CARRIERS, for which data were available on receptor status (Supplementary Table S5, available at https://doi.org/10.1016/j.annonc.2024.07.244), stronger associations were evident when analysis was restricted to just oestrogen receptor (ER)-negative breast cancers (OR 3.18, 95% CI 1.99-5.09 for *RAD51C*; OR 3.21, 95% CI 1.83-5.65 for *RAD51D*; OR 4.41, 95% CI 2.87-6.78 for *BARD1*). Association metrics were further strengthened by restricting the analysis to just triple-negative breast cancer cases (OR 4.32, 95% CI 2.35-7.94 for *RAD51C*; OR 5.05, 95% CI 2.42-10.53 for *RAD51D*; OR 6.26, 95% CI 3.57-10.99 for *BARD1*). Notably, the frequency of PVs in 'population-type' breast cancer cases across the three datasets was very low for *RAD51C* (0.11%, or 1 in 913), for *RAD51D* (0.09%, or 1 in 1079) and for *BARD1* (0.15%, or 1 in 672), with higher but still modest PV frequency in triple-negative breast cancers of 1 in 307, 1 in 430 and 1 in 239, respectively.

Also noteworthy is the markedly stronger association with triple-negative breast cancer compared with 'population-type' breast cancer. *BRCA1* showed an OR of 52.23 (95% CI 39.90-68.38) versus 8.96 (95% CI 7.67-10.48), while *PALB2* exhibited an OR of 11.31 (95% CI 7.77-16.47) versus 4.26 (95% CI 3.64-4.99).

Conversely, compared with the association for all breast cancers, association with ER-negative and triple-negative disease was substantially lower for *CHEK2* and nonsignificant for *ATM*: these genes are evidently predominantly associated with ER-positive disease.

Genes associated with pleiomorphic cancer syndromes: TP53, PTEN, NF1, STK11 and CDH1

For *TP53*, PV counts were low across the studies. For BRIDGES, the published association for PTVs was nonsignificant, with 7/48 826 PTVs in cases versus 2/50 703 in controls producing an OR of 3.06 (95% CI 0.63-14.91; Supplementary Table S2A and B, available at https://doi.org/10.1016/j.annonc.2024.07.244). Adding in the missense PVs increased these numbers to 34/48 826 in cases versus 9/50 703 (0.018%) in controls, producing an OR of 3.30 (95% CI 1.58-6.89). In CARRIERS, the reported frequency for *TP53* of all PVs was 19/32 247 in cases and 2/32 544 (0.0061%) in controls; although an OR was not presented due to small numbers, these numbers would correspond to an OR of 9.59 (95% CI 2.23-41.19). In UKB, for *TP53* the frequency of all PVs was 2/20 324 in breast cancer cases and 12/229 697 (0.0052%) in controls, equating to an OR of 1.88 (95% CI 0.42-8.45). The observed frequency of PVs in controls was similar between UKB and CARRIERS but higher in BRIDGES; these differences in control frequencies likely reflect technical differences between studies, with aggressive filtering for low VAF variants of putative haematopoietic clonal origin applied in CARRIERS and UKB but not in BRIDGES. We calculated the overall weighted average OR for all PVs across the three studies to be 3.62 (95% CI 1.98-6.61), with a BC PV frequency of 0.054% (or 1 in 1844) and a control PV frequency of 0.01% (1 in 13 606).

For each of *PTEN* and *NF1*, PV frequencies and metrics of association were consistent across all studies. The weighted average for all PVs yielded associations as follows: OR 2.01 (95% CI 1.43-2.83) for *NF1* and OR 2.63 (95% CI 1.38-5.02) for *PTEN.* BC PV frequencies were 0.068% (1 in 1470) for *NF1* and 0.027% (1 in 3755) for *PTEN.*

For *STK11*, the published results in BRIDGES for PTVs were nonsignificant (OR 1.60, 95% CI 0.48-5.28) with observation of PVs in 6/48 826 cases. *STK11* was not investigated in the CARRIERS study. In UKB, there were 0/20 324 observations of PVs in cases and only 1/229 697 in controls. This analysis produced a weighted average OR for all PVs combining BRIDGES and UKB of 1.60 (95% CI 0.48-5.30) with a BC PV frequency of 0.0087% (1 in 11 525).

For *CDH1*, the published results showed borderline significance in CARRIERS with an OR of 2.50 (95% CI 1.01-7.07) for all PVs, while nonsignificant findings were observed in BRIDGES with an OR of 0.86 (95% CI 0.37-1.98) for PTVs alone. In UKB, the estimate was higher at an OR of 4.11 (95% CI 1.92-8.81) for all PVs. Incorporating these findings with additional non-PTV PVs for BRIDGES yielded a weighted average OR of 2.01 (95% CI 1.25-3.24), with a BC PV frequency of 0.037% (1 in 2668). When restricting our analysis to breast cancer cases with lobular histology, the OR increased to 22.01 (95% CI 9.45-51.31), with PVs identified in 14 out of 6107 cases (0.23%, 1 in 436) of confirmed lobular breast cancers. These cases comprised 15.3% of all breast cancers in UKB and 9.3% in CARRIERS[12,13] (Supplementary Table S7, available at https://doi.org/10.

1016/j.annonc.2024.07.244). There was no evidence of association between *CDH1* and breast cancer of nonlobular/unknown histology (OR 1.09, 95% CI 0.26-4.59).

Mismatch repair genes (MLH1, MSH2, MSH6, PMS2 and EPCAM) and other putative breast cancer susceptibility genes

In the weighted meta-analysis, there was no significant association between breast cancer and any of the mismatch repair genes. Notably, while an association for *MSH6* had been previously reported and observed in the BRIDGES study (OR 1.96, 95% CI 1.15-3.33), the weighted meta-analysis showed a nonsignificant association (OR 1.27, 95% CI 0.98-1.66).[14] *PMS2* was not included in the CARRIERS study. Although no association was observed in either study, there was significant heterogeneity in variant counts between BRIDGES and UKB, potentially reflecting technical differences in mapping (see Table 1, Supplementary Table S1, available at https://doi.org/10.1016/j.annonc.2024.07.244).

CDKN2A was recently reported as associated with breast cancer by Wilcox et al.[15] in a gene discovery analysis combining 8238 enriched (familial, young-onset and bilateral) breast cancer cases with 17 958 UKB breast cancer cases. This association was primarily driven by UKB but was nonsignificant in the enriched series. *CDKN2A* was not reported in the BRIDGES analysis and was nonsignificant in the CARRIERS analysis (OR 1.51, 95% CI 0.47-5.22). We reproduced the previously reported association in UKB (OR 1.98, 95% CI 1.23-3.20). Upon meta-analysis with CARRIERS, the association signal from UKB was attenuated but persisted (OR 1.91, 95% CI 1.22-2.98). Notably, the control frequency of PVs is approximately threefold higher in UKB than in CARRIERS (Supplementary Table S4, available at https://doi.org/10.1016/j.annonc.2024.07.244).

Association metrics (ORs) were nonsignificant on weighted meta-analysis across the three studies for other previously proposed putative BCSGs included in the BRIDGES and/or CARRIERS analyses, including *ABRAXAS1, AKT1, BABAM2, NBN, PIK3CA, RAD50, RECQL, RINT1, SLX4* and *XRCC2* (Supplementary Table S2A and B, available at https://doi.org/10.1016/j.annonc.2024.07.244).[16-20]

DISCUSSION

We present data from three population-ascertained studies of breast cancer cases and controls, examining how the frequencies of PVs and accordant ORs of association for 37 putative BCSGs vary across the studies. We explored the differences in patient inclusion and molecular and analytical methodologies between the studies (Supplementary Table S1, available at https://doi.org/10.1016/j.annonc.2024.07.244). However, as a result of variant rarity and per-gene study-specific vagaries underpinning observed differences, although there is heterogeneity, we propose high utility afforded in the integration of data from these three studies to generate unified, best-powered estimates for BCSG PV frequencies and associations. We thus present weighted average breast cancer ORs and PV frequencies comprising in total 101 397 breast cancer cases and 312 944 controls. We propose these metrics will be of significant utility for (i) clinical management guidance and (ii) downstream public health-related economic or capacity analyses. These integrated summary estimates from unselected, population-ascertained breast cancer cases are particularly informative in the context of the expansion of BCSG germline testing to an increasingly wide group of breast cancer cases. For example, in the recent guidelines from the American Society of Clinical Oncology (ASCO), germline genetic testing was recommended for all women with a past or current diagnosis of breast cancer at age $\leq$65 years.[21]

Notably, these integrated metrics provide unequivocal delineation of prevalence-risk impact for *BRCA1, BRCA2* and *PALB2,* with a BC PV frequency of 3% (combined) in 'population-type' breast cancer cases. These integrated data also clearly illustrate the extremely low BC PV frequency for the 'syndromic' genes: 1 in 1470 for *NF1*, 1 in 1844 for *TP53*, 1 in 3755 for *PTEN*, 1 in 2668 for *CDH1* and 1 in 11 525 for *STK11*, with PVs for these genes occurring in 'population-type' breast cancer at rates only modestly higher than in the background population. The BC PV frequency was also low for *BARD1* (1 in 672), *RAD51C* (1 in 913) and *RAD51D* (1 in 1079). Furthermore, the increased data volume enables clearer delineation of breast cancer subtype specificity; for example, these data would support the association of *CDH1* being specifically with lobular breast cancer (OR 22.01, 95% CI 9.45-51.31), given the clear absence of signal in breast cancers of nonlobular/unknown pathology (OR 1.09, 95% CI 0.26-4.59).

Interstudy heterogeneity and limitations regarding data integration

Notably, for all of the studies, only small variants within or close to exons were included in the analyses, meaning copy number and deep intronic PVs were not counted in the total number of observed PVs. This undercount on PV frequency will vary according to the per-gene spectrum of mutational mechanism but would be anticipated to be nondifferential between cases and controls, and thus not be anticipated to impact on the observed metrics of association.

In the context of analysing 37 genes, many of which have low PV frequency, it is not straightforward to discern whether heterogeneity observed between studies is influenced by the differences in case inclusion, laboratory or analytical methodology or is due to statistical chance. Notably, the pattern of differences in PV frequency between the three studies varies across the 37 genes.

Technical factors relating to sequencing, quality control (QC) and variant inclusion would be anticipated to influence PV frequency; however, these are expected to be non-differential between cases and controls within the given study (i.e. not distorting the per-study OR). For example, the differences in the molecular genetic methodologies and probe selections underlying the various panel and exome

capture kits potentially may introduce between-study differences in coverage and mappability. Whole exome/genome sequencing undertaken for UKB was of lower mean coverage than the custom panels used for BRIDGES and CARRIERS, potentially resulting in under-ascertainment of PVs within UKB in regions with poorer callability (see Supplementary Methods, available at https://doi.org/10.1016/j.annonc.2024.07.244). Inclusion and classification as pathogenic of the non-PTV variants also differed between studies. These technical differences likely explain some of the variation observed between the three studies. For example, the lower control frequency of *BRCA1* PVs in UKB compared with BRIDGES (with CARRIERS midway between) may reflect the lower coverage of this gene in UKB. For *CHEK2*, CARRIERS exhibited a significantly lower frequency of all PVs compared with UKB and BRIDGES, perhaps reflecting the capture. For *PMS2*, the lower frequency of PVs in BRIDGES compared with UKB may reflect mapping issues related to known pseudogenes. For *TP53* and *NF1*, the stricter VAF threshold used for CARRIERS and UKB may account for the comparatively higher PV frequency observed in BRIDGES, likely reflecting a proportion of observed variants being artefacts of clonal haematopoiesis of indeterminate potential (see Supplementary Table S1, available at https://doi.org/10.1016/j.annonc.2024.07.244).

There were some modest differences between the three constituent studies regarding the populations of breast cancer cases (age of diagnosis, proportion of noninvasive cancers, histology and receptor status; Supplementary Tables S1 and S5, available at https://doi.org/10.1016/j.annonc.2024.07.244). Given the subtype-specific gene associations, these differences might, for relevant genes, result in enrichment among cases for PVs, thus (modestly) influencing association metrics. Notably, receptor status was unavailable for UKB cases, with substantial data missingness also noted for BRIDGES and CARRIERS (Supplementary Table S6, available at https://doi.org/10.1016/j.annonc.2024.07.244).

All three constituent studies were majority white/European with CARRIERS including more samples from extremely diverse ancestry populations (Supplementary Table S1, available at https://doi.org/10.1016/j.annonc.2024.07.244). It is unlikely that the different ancestry composition of the three studies will impact PV frequency (in the absence of sizeable presence of founder populations) nor the magnitude of association (ethnicity/country being included as covariates in regressions to avoid confounding by population stratification).

Summary metrics were used in the meta-analysis for BRIDGES and CARRIERS, meaning that we were unable to align upstream quality control processes, regression methodology and variant inclusion between these studies. Analyses integrating individual-level data across the three studies would provide an opportunity to apply standardised QC thresholds and variant classification, as well as to provide ORs stratified by age and additional factors. Nevertheless, due to the inherent heterogeneity of the contributing data in capture design and sequencing depth, standardised QC would potentially result in a sizeable loss of data. Summary outcome metrics would still necessarily be predicated upon the pre-existing characteristics of the constituent studies, with regard to both participant ascertainment and molecular analyses.

Implications for breast cancer risk in a clinical setting

Easton and colleagues[22] in 2015 articulated for the first time what have since widely become established definitions for a high-penetrance BCSG (OR/relative risk (RR) $\geq$4) and a moderate-penetrance BCSG (OR/RR $\geq$2). Evaluating published studies, the 2015 group asserted *BRCA1*, *BRCA2* and *TP53* as being definite high-penetrance BCSGs while *PALB2*, *PTEN*, *CDH1*, *STK11*, *NF1*, *ATM*, *CHEK2* and *NBN* were deemed to be 'likely' BCSGs of at least moderate penetrance. In their analysis, the group focused on the lower 90% CI and the point estimate of the association metric (RR/OR). The group at the time highlighted the lack of availability of unbiased data quantifying breast cancer risk, particularly for the 'syndromic' genes (e.g. the breast cancer risk estimate for *TP53* cited as best available in 2015 was RR 105, 90% CI 62-165, derived from family studies).[23]

Revisiting these assignations nearly a decade on, with these meta-analysed data for >100 000 'population-type' breast cancer cases and >300 000 controls, genes meeting the RR/OR $\geq$4 point estimate threshold for high penetrance were only *BRCA1*, *BRCA2* and *PALB2* (although for *PALB2*, the lower 90% CI sits just below 4: OR 4.30, 90% CI 3.78-4.90; Supplementary Table S4, available at https://doi.org/10.1016/j.annonc.2024.07.244). Restricting by breast cancer subtype, *CDH1* would qualify as high penetrance for lobular breast cancer. Nonsyndromic genes meeting the RR/OR $\geq$2 point estimate threshold for moderate penetrance were *CHEK2*, *ATM* and *BARD1* (although again the lower 90% CIs are below 2 for *ATM* and *BARD1*: *ATM* OR 2.16, 90% CI 1.97-2.37 and *BARD1* OR 2.34, 90% CI 1.92-2.86). *RAD51C*, *RAD51D* and *BARD1* would qualify as being of moderate or high penetrance if restricting the analysis to triple-negative breast cancer (depending on the use of lower 90% CI or point estimates).

The metrics of association (ORs) from the meta-analysed data for the 'syndromic' genes are strikingly modest. Part of the motivation for conducting studies of 'unselected' or 'population-type' breast cancers, such as BRIDGES and CARRIERS, was to obviate the overestimated risks from ascertainment bias inherent to previous family-based segregation studies. However, the routine clinic-based case ascertainment for BRIDGES and CARRIERS and in particular the national cohort study recruitment for UKB may potentially result in depletion for those breast cancer cases presenting younger (especially if lethal) and women whose breast cancer is accompanied by additional phenotypes (e.g. intellectual disability or multiple early-onset cancers). Therefore, the ORs calculated in these population-based studies may potentially *underestimate*

the true breast cancer risks, particularly for these genes in which variation in age-stratified risk is likely to be particularly marked.[24] However, the very low PV frequencies for these genes in these meta-analysed data are nonetheless highly informative in the context of germline testing in mainstream breast cancer clinics (where very young-onset, 'syndromic' or complex multicase familial disease would likely have been filtered out).

CONCLUSIONS

We have presented three population-based breast cancer case—control analyses and weighted meta-analysis comprising 101 397 women with breast cancer and 312 944 control women, considering how composition of patients, analytical methodology, PV inclusion and chance variation in rare variant numbers may have contributed to the variation observed between studies for PV frequencies and associations for the 37 putative BCSGs analysed. Many current breast cancer panels include a large number of genes. This meta-analysis demonstrates that, for many of these genes, the PV frequency in 'population-type' breast cancer is very low, and the magnitude of association is modest. Germline genetic testing is being recommended for an increasingly wide group of women with breast cancer. These data will potentially be informative for delineating the optimal gene set for germline testing in this 'population-type' group of breast cancers, to ensure net clinical benefit over harm to the individual patient and the healthcare system.

FUNDING

This work was supported by CG-MAVE, Cancer Research UK Programme Award [grant number EDDPGM-Nov22/100004 to CT funding CFR and SA] and the Canadian Institutes of Health Research [grant number FDN-148390 to WDF]. MT was supported by the NIHR Cambridge Biomedical Research Centre (grant number NIHR203312). HH was supported by the NIHR Exeter Biomedical Research Centre (grant number NIHR203320).

DISCLOSURE

JB reports receipt of honoraria for participation in an AstraZeneca Steering Committee and AstraZeneca educational programs; receipt of honoraria for being an invited speaker from AstraZeneca and receipt of honoraria for role as local PI from MEDSIR. SMD reports receipt of honoraria for acting as a consultant from GSK and Intellia; is an ASCO member with no financial interest and is a local PI for AstraZeneca with no financial interest. DGE reports receipt of honoraria for consultancy on genetic testing for PARPi from AstraZeneca and receipt of honoraria for involvement on advisory boards from Everything Genetic Ltd, Recursion, Springworks and Syantra. HH reports receipt of honoraria for participation on an advisory board from AstraZeneca and nonfinancial interest for a leadership role as Chair of the UK Cancer Genetics Group (UKCGG). NH is the Chair of the European Reference Network Genetic Tumour Risk Syndromes (ERN GENTURIS). MR reports receipt of honoraria for review of guideline pathways from Change Healthcare; receipt of honoraria for being an invited speaker at CME events from Clinical Care Options, MJM Holdings, and Physician's Education Resource; receipt of honoraria for participation in an advisory board from myMedEd; receipt of honoraria for role as local PI from Artios Pharma and Merck; receipt of funding for the research study ICEBERG from AstraZeneca, participation in AstraZeneca Steering Committee with no financial interest or personal compensation; role as co-PI for Merck IIT of neoadjuvant olaparib/pembrolizumab in BRCA carriers; participation in Merck Steering Committee with no financial compensation; role as co-PI for Pfizer of clinical trial of ZEN03694 and talazoparib in TNBC; advisory roles without financial interest for Foundation Medicine, Tempus Labs, and Zenith Pharmaceuticals; editorial service for medical writing of reports resulting from the OlympiAD trial with no financial interest from AstraZeneca; editorial services for write-ups of trials with no financial interest from Merck; and is an ASCO member with no financial interest. WDF reports nonfinancial interest in an HBOC leadership role. CT reports honoraria paid to charity as an invited speaker from AstraZeneca and honoraria paid to charity for participation in an advisory board for an observational study from Roche. All remaining authors have declared no conflicts of interest.

REFERENCES

1. 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.
2. Wooster R, Bignell G, Lancaster J, et al. Identification of the breast cancer susceptibility gene BRCA2. *Nature*. 1995;378(6559):789-792.
3. Loveday C, Turnbull C, Ramsay E, et al. Germline mutations in RAD51D confer susceptibility to ovarian cancer. *Nat Genet*. 2011;43(9):879-882.
4. Loveday C, Turnbull C, Ruark E, et al. Germline RAD51C mutations confer susceptibility to ovarian cancer. *Nat Genet*. 2012;44(5):475-476. author reply 476.
5. 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.
6. Erkko H, Xia B, Nikkila J, et al. A recurrent mutation in PALB2 in Finnish cancer families. *Nature*. 2007;446(7133):316-319.
7. Renwick A, Thompson D, Seal S, et al. ATM mutations that cause ataxia-telangiectasia are breast cancer susceptibility alleles. *Nat Genet*. 2006;38(8):873-875.
8. 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.
9. 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.
10. Hu C, Hart SN, Gnanaolivu R, et al. A population-based study of genes previously implicated in breast cancer. *N Engl J Med*. 2021;384(5):440-451.
11. Backman JD, Li AH, Marcketta A, et al. Exome sequencing and analysis of 454,787 UK Biobank participants. *Nature*. 2021;599(7886):628-634.
12. Yadav S, Hu C, Nathanson KL, et al. Germline pathogenic variants in cancer predisposition genes among women with invasive lobular carcinoma of the breast. *J Clin Oncol*. 2021;39(35):3918-3926.
13. van Veen EM, Evans DG, Harkness EF, et al. Extended gene panel testing in lobular breast cancer. *Fam Cancer*. 2022;21(2):129-136.
14. Couch FJ, Shimelis H, Hu C, et al. Associations between cancer predisposition testing panel genes and breast cancer. *JAMA Oncol*. 2017;3(9):1190-1196.

15. Wilcox N, Dumont M, González-Neira A, et al. Exome sequencing identifies breast cancer susceptibility genes and defines the contribution of coding variants to breast cancer risk. *Nat Genet*. 2023;55(9):1435-1439.
16. Heikkinen K, Rapakko K, Karppinen SM, et al. RAD50 and NBS1 are breast cancer susceptibility genes associated with genomic instability. *Carcinogenesis*. 2006;27(8):1593-1599.
17. Cybulski C, Carrot-Zhang J, Kluźniak W, et al. Germline RECQL mutations are associated with breast cancer susceptibility. *Nat Genet*. 2015;47(6):643-646.
18. Tommiska J, Seal S, Renwick A, et al. Evaluation of RAD50 in familial breast cancer predisposition. *Int J Cancer.* 2006;118(11):2911-2916.
19. Landwehr R, Bogdanova NV, Antonenkova N, et al. Mutation analysis of the SLX4/FANCP gene in hereditary breast cancer. *Breast Cancer Res Treat*. 2011;130(3):1021-1028.
20. Seal S, Thompson D, Renwick A, et al. Truncating mutations in the Fanconi anemia J gene BRIP1 are low-penetrance breast cancer susceptibility alleles. *Nature Genet*. 2006;38(11):1239-1241.
21. Bedrosian IA-O, Somerfield MA-O, Achatz MI, et al. Germline testing in patients with breast cancer: ASCO-Society of Surgical Oncology Guideline. *J Clin Oncol*. 2024;42:584-604.
22. Easton DF, Pharoah PD, Antoniou AC, et al. Gene-panel sequencing and the prediction of breast-cancer risk. *N Eng J Med*. 2015;372(23):2243-2257.
23. Hwang SJ, Lozano G, Amos CI, Strong LC. Germline p53 mutations in a cohort with childhood sarcoma: sex differences in cancer risk. *Am J Hum Genet*. 2003;72(4):975-983.
24. Evans DG, Howell SJ, Frayling IM, Peltonen J. Gene panel testing for breast cancer should not be used to confirm syndromic gene associations. *NPJ Genom Med*. 2018;3:32.