Supplementary Appendix

Table of Contents

Supplementary Methods

Enrichment protocols used for constituent studies

Our three selected datasets (BRIDGES, CARRIERS and UK Biobank) employed distinct enrichment methodologies for amplification of target regions. In BRIDGES, sequencing libraries were constructed using the Fluidigm Juno platform, in which target regions are amplified via assay primers, uniquely barcoded and subsequently sequenced using an Illumina HiSeq 4000. The CARRIERS sequencing library was constructed using a QIAseq Targeted DNA Panel, in which targets are enriched using gene-specific reverse primers and universal forward primers; libraries were then also sequenced on an Illumina HiSeq 4000. The UK Biobank exome data were generated using the IDT xGen Exome Research Panel, which uses individually synthesised biotin-modified oligonucleotide probes to amplify target regions. The resulting library was then sequenced on the Illumina NovaSeq 6000 platform.

Differences in primer design, reaction chemistry and sequencing platform may naturally lead to biases in amplification and read quality; this may be particularly marked for genes harbouring low complexity regions or for which primers in flanking regions may exhibit off-target binding.

The use of the aggregated population VCF in the UKB arm of analysis did not allow quantification of gene-level coverage due to absence of data at non-variant positions; there may thus be slight under-ascertainment of pathogenic variants in regions with poorer callability compared to the higher-coverage panel workflows used in BRIDGES and CARRIERS

Curation of UK Biobank breast cancer and control cohorts

For each individual in the UK Biobank dataset, all possible instances were extracted for the following data fields:

- Self-reported sex (p31)
- Genetic sex (p22011)
- Age (p21022)
- Self-reported ethnicity (p21000_i0-2)
- Sex chromosome aneuploidy (p22019)
- Genetic relatedness pairing (p22011_a0-4)
- Genetic relatedness factor (p22012_a0-4)
- Self-reported cancer code (p20001_i0-3)
- Type of cancer: ICD9 (p40013_i0-14)
- Type of cancer: ICD10 (p40006_i0-21)
- Histology of cancer tumour (p40011_i0-21)

The total cohort was filtered to retain only individuals with both a self-declared (p31) and genetic (p22011) sex of “Female”, provided there was no evidence of sex chromosome aneuploidy (p22019) and the participant had not withdrawn from the study.

Individuals with any recorded instances of the following ICD10 (p40006) and ICD9 (p40013) codes indicating breast cancer in cancer registry-linked UKB fields were assigned to the case cohort, regardless of whether breast cancer was the first reported cancer type for the individual:

Category	ICD10	ICD9
Malignant neoplasm of breast	C50.X	174X
Carcinoma *in situ* of breast	D05.X	2330
Neoplasm of uncertain behaviour of breast	D48.6	2383

We further included in the case cohort any individuals with self-reported cancer type (p20001) including the term “breast”. Any individual failing to meet these criteria was assumed to not have had breast cancer and was assigned to the control cohort.

Related individuals, defined as pairs of individuals with a genetic relatedness factor (p22012) ≥0.17, were filtered to retain a single individual from each pair, preferentially retaining individuals from the case cohort where a pair consisted of one individual from each of the case and control cohorts, otherwise randomly selecting an individual to retain.

Granular ethnicities (p21000) were collapsed according to their top-level parent ethnicity (one of “White”, “Black”, “Asian”, “Chinese” or “Mixed”). To avoid confounding downstream logistic regression, individuals were excluded from either cohort if they were associated with multiple ethnicities with discrepant parent terms or where their self-reported ethnicity belonged to the “Other” parent term.

For downstream sub-analyses of lobular carcinoma ORs, the case cohort was partitioned according to the content of the cancer histology field (p40011), with lobular status being ascribed to cases with any of the following annotations:

- Lobular carcinoma, NOS
- Infiltrating duct or lobular carcinoma
- Infiltrating lobular mixed with other types of carcinoma

Cases were otherwise ascribed “non-lobular” status. Any individual with any recorded instance of lobular carcinoma was assigned to the lobular cohort, regardless of whether they had any additional non-lobular breast cancer histology record.

Selection of genes and transcripts of interest and variant extraction from UK Biobank

37 genes were selected for analysis, encompassing all genes investigated in one or both of BRIDGES and CARRIERS, but with the exclusion of *MUTYH* and *PPM1D*. Comparison of the RefSeq transcripts selected for inclusion in those constituent studies demonstrated near-complete concordance, with only *NF1* and *RAD51D* differing between BRIDGES and CARRIERS; in both cases, the MANE Select transcript (NM_001042492.3 and NM_002878.4, respectively) was selected for downstream analyses.

For all genes except one, variants were annotated against the respective selected RefSeq transcript; however, for *CDKN2A*, the Ensembl transcript ENST00000304494.5 was used, as it contained a longer 5' UTR encompassing a known pathogenic *CDKN2A* variant. This variant was annotated as an upstream gene variant when using its canonical RefSeq transcript (NM_000077.5).

BED files were generated containing the co-ordinates of all exons in selected transcripts of interest ± 25 intronic bases. Variants overlapping these regions in the corresponding UK Biobank population VCF files were extracted using the BCFtools view command via the UKB DNAnexus platform.

Gene	Transcript for annotation
ABRAXAS1	NM_139076.3
AKT1	NM_005163.2
ATM	NM_000051.4
BABAM2	NM_004899.5
BARD1	NM_000465.4
BLM	NM_000057.4
BRCA1	NM_007294.4
BRCA2	NM_000059.4
BRIP1	NM_032043.3
CDH1	NM_004360.5
CDKN2A	ENST00000304494.5
CHEK2	NM_007194.4
EPCAM	NM_002354.3
ERCC3	NM_000122.2
FANCC	NM_000136.3
FANCM	NM_020937.4
GEN1	NM_182625.5
MEN1	NM_130799.2
MLH1	NM_000249.4
MRE11	NM_005591.4
MSH2	NM_000251.3
MSH6	NM_000179.3
NBN	NM_002485.5
NF1	NM_001042492.3
PALB2	NM_024675.4

PIK3CA	NM_006218.4
PMS2	NM_000535.7
PTEN	NM_000314.8
RAD50	NM_005732.4
RAD51C	NM_058216.3
RAD51D	NM_002878.4
RECQL	NM_002907.4
RINT1	NM_021930.6
SLX4	NM_032444.4
STK11	NM_000455.5
TP53	NM_000546.6
XRCC2	NM_005431.2

Generation of pathogenic variant cohorts

Variants in genes of interest identified from the UKB data were annotated using Ensembl VEP (v110). We ascribed as pathogenic all putatively protein-truncating variants, namely any variants with an annotation of stop_gained, frameshift_variant, splice_acceptor_variant or splice_donor_variant that lay upstream of the final exon. In accordance with the BRIDGES methodology, assignation of pathogenicity to final exon PTVs was permitted for 6 genes with prior evidence of C-terminal essentiality (*ATM*, *BARD1*, *BRCA1*, *PALB2*, *RAD51C* and *RAD51D*). We additionally excluded from assignation of pathogenicity specific *BRCA1* canonical splice variants classified as uncertain significance according to ENIGMA guidelines, or impacting residues within tandem acceptor sites. We excluded one *FANCM* variant, c.1397-3_1397-2del, from the pathogenic variant set as it was present at too high a frequency to be pathogenic. These variants formed the “PTV” tier of putative pathogenic variants.

A second tier of pathogenic variants was generated using the May 2023 ClinVar flat file; pathogenicity was ascribed to variants observed in UKB if they were classified as pathogenic or likely pathogenic in ClinVar with ≥2* review status and no annotation of ‘risk allele’ status or conflicting interpretations. These variants formed the “non-PTV-PV” tier of variants.

Generation of betas for breast cancer association via logistic regression

Betas of association with breast cancer status were generated via logistic regression using bespoke R scripts. Taking breast cancer status as the response variable, the regression model calculated the strength of association between mutation status of an individual for each gene and tier of variant (PTVs, non-PTV-PVs and “all PVs”, i.e. the presence of any PTV and/or any non-PTV-PV), adjusting for participant top-level ethnicity and age at recruitment.

Generation of BRIDGES pathogenic non-PTV-PV variant cohort

Non-PTV-PVs were incorporated into both the CARRIERS analysis and our UKB-based approach. To ensure variant sets were consistent between constituent datasets in our meta-analysis, we supplemented the BRIDGES cohort of PTVs with pathogenic missense variants available for download from the original BRIDGES publication as Supplementary Material. No other variant consequences besides missense were available to download.

The genomic co-ordinates of missense variants identified in the population-based constituent studies of BRIDGES were cross-referenced with the May 2023 ClinVar release flat file as above to identify variants with existing clinical classifications. As with variants observed in UKB, BRIDGES missense variants were considered pathogenic only in the presence of non-conflicting ClinVar classifications of “Pathogenic”, “Likely pathogenic” or “Pathogenic/Likely_pathogenic” with ≥ 2* review status. We then calculated unadjusted betas for the association of these variants with breast cancer case status using the raw variant counts in cases and controls.

As no participant-level metadata was available to allow adjustment of missense variant betas by substudy or ethnicity, we used the betas calculated for the BRIDGES PTV data to generate gene-specific beta adjustment factors. These were calculated as the arithmetic difference between the published adjusted beta (generated by logistic regression) and the unadjusted beta calculated from the raw variant counts for a given gene. We observed that, for most surveyed genes, betas remained minimally affected by logistic regression adjustment; further, little to no difference was observed in the standard error of the odds ratio estimate (Supplementary Figure 3). This also held true for unadjusted and adjusted betas in CARRIERS and our UKB analysis.

BRIDGES missense betas were then adjusted by the addition of their corresponding adjustment factor.

Methodology for meta-analysis and calculation of associated statistics

The BRIDGES and CARRIERS studies provided summary-level data for each surveyed gene in the form of odds ratios (ORs) and 95% confidence intervals (CIs). To integrate the findings from BRIDGES and CARRIERS into our meta-analysis, it was necessary to estimate standard error (SE) from the published summary-level values alone.

95% CIs for a given beta (ln(OR)) can be calculated using the following formula:

$$95\%\ CI = e^{\ln(OR) \pm (1.96 * SE)}$$

Through rearrangement of this equation, we calculated two estimates for SE for each gene and study (one each based on the upper and lower CI limits) as follows:

$$SE = \frac{\ln(OR) - \ln(Lower\ 95\%\ CI)}{1.96} = \frac{\ln(\frac{OR}{Lower\ 95\%\ CI})}{1.96}$$

$$SE = \frac{\ln(Upper\ 95\%\ CI) - \ln(OR)}{1.96} = \frac{\ln(\frac{Upper\ 95\%\ CI}{OR})}{1.96}$$

To account for the low precision given in the summary data from both BRIDGES and CARRIERS, we took as our final SE estimate for a given gene and study the mean average of the two estimates.

Meta-analyses were conducted using a fixed effect inverse-variance approach, weighting effect sizes from the constituent studies by the reciprocal of the square of their standard error (SE) as follows:

$$Weighted\ \beta = \frac{\sum w_i \cdot \beta_i}{\sum w_i}$$

Where β_i is the estimated beta of association in constituent study i, and w_i is the weight of the respective study, defined as $\frac{1}{SE^2}$

A combined estimate of standard error was calculated using the formula:

$$Weighted\ SE = \sqrt{(\sum w_i)^{-1}}$$

With w_i defined as above.

The upper and lower 95% confidence limits of the weighted ORs were then calculated using the above equation. We further generated 90% CIs by replacing 1.96 with 1.645 for the coefficient of the standard error in the CI calculation.

Our primary meta-analyses comprised combination of the BRIDGES PTV and pathogenic missense datasets into an overall BRIDGES pathogenic dataset, and the meta-analysis of that resulting dataset, CARRIERS and the UKB analysis. We further meta-analysed the two datasets from BRIDGES and CARRIERS illustrating OR by breast cancer receptor status (ER-positive, ER-negative and triple-negative) for a subset of genes of interest.

For each gene, we additionally calculated two primary metrics of inter-study heterogeneity; namely, the Cochran's Q statistic, and the related I^2 metric. Cochran's Q was calculated as follows:

$$Q = \sum w_i \, (\beta_i - \frac{\sum w_i \cdot \beta_i}{\sum w_i})^2$$

Where w_i and β_i are the weight and estimated beta in each constituent study, as above (NB: $\frac{\sum w_i \cdot \beta_i}{\sum w_i}$ is equivalent to the combined beta for a given gene). p-values were generated by referring the Q statistic to a χ^2 distribution with n-1 degrees of freedom (with n=number of constituent studies).

As Cochran's Q may be biased against meta-analyses comprising low numbers of studies, such as the analyses presented here, we further computed the I^2 statistic as follows:

$$I^2 = \frac{Q - (n-1)}{Q}$$

Where Q = Cochran's Q, as derived above, and n-1 = the degrees of freedom from n constituent studies.

Supplementary Tables

Supplementary Table 1: Descriptive summary of the BRIDGES, CARRIERS, and UK Biobank data. BC = breast cancer, TNT = triple-negative tumour, SD = standard deviation, PTV = protein-truncating variant, PV = pathogenic variant, P/LP = pathogenic/likely pathogenic. Age ranges represent age at diagnosis for cases and age at enrolment for controls. *For BRIDGES, the invasiveness status is only reported for the full series of participants (60,466 breast cancer cases, which includes in addition to the 48,826 cases from population-based studies, 11640 cases from studies of selective recruitment). **Histology information for CARRIERS cohort obtained from Yadav et al., 2021[1].

		BRIDGES		CARRIERS		UKBIOBANK	
		BC Cases	Controls	BC Cases	Controls	BC Cases	Controls
Number		48,826	50,703	32,247	32,544	20,324	229,697
Mean age (SD)		Median of constituent study mean ages: 56.55	Median of constituent study mean ages: 56.3	60.44 (11.85)	60.66 (11.81)	57.8 (9.9)	56.2 (8.0)
Age range		42.1-68.4	33.2-66.7	21.00-94.00	21.80-94.30	18-82	39-71
Source Studies		30 population- and hospital-based studies participating in the BCAC; cases unselected for family history		12 American population-based studies		Females identified from UK Biobank with record of BC diagnosis for cases, or lack thereof for controls	
Distribution of breast cancer subtypes	**Invasiveness***	2.7% status missing where status available (n=58,811): 92.9% invasive / 7.1% DCIS		3.2% status missing where status available (n=31,221): 85.8% invasive / 14.2% DCIS		2.8% status missing where status available (n=19,746): 87.1% invasive/12.9% in situ (inc. lobular in situ)	
	ER status	21.7% status missing where status available (n=38,232): 79.7% ER pos / 20.3% ER neg		31.1% status missing where status available (n=22,233): 82.9% ER pos / 17.1% ER neg		100% status missing	
	TNT status	26.5% status missing where status available (n=35,863): 7.9% TNT / 92.1% non-TNT		59.9% status missing where status available (n=12,915): 11.3% TNT / 88.7% non-TNT		100% status missing	
	Histology**	100% status missing		27.7% status missing/other where status available (n=23,322): 12.9% lobular / 87.1% ductal		13.8% status missing/other where status available (n=17,512): 18.2% lobular/71.8% ductal	
Ancestry		Individual-level ancestry not explicitly presented 27 studies from Europe, North America, Australasia (~88.7% of samples post QC) 2 studies from Asia (~10.2% of samples post QC 1 study form Columbia (~1.1% of samples post QC)		Non-Hispanic White (74.3%) African American (13.4%) Hispanic (5.6%) Asian (4.9%) Other (1.7%)	Non-Hispanic White (75.8%) African American (14.6%) Hispanic (4.1%) Asian (3.9%) Other (1.7%)	White (97.0%) Asian (1.3%) Black (1.0%) Mixed (0.5%) Chinese (0.2%)	White (95.5%) Asian (1.7%) Black (1.7%) Mixed (0.7%) Chinese (0.4%)
Panel/sequencing		Custom 34-gene Fluidigm Juno panel sequenced on Illumina HiSeq 4000, achieving mean coverage of 349 reads; 91.1% of target positions callable (≥15x coverage and base quality ≥20)		Custom 37-gene QIAseq panel sequenced on Illumina HiSeq 4000; read depth ≥20 reads for 99.3% of target regions		Exome Research Panel (v1.0, IDT xGen targeting 39 Mbp of the human genome) sequenced on Illumina NovaSeq 6000; upstream QC and processing using the DNAnexus OQFE protocol (CITE PMID: 30840781)	
Quality Control (QC)		Sample QC (kinship filtering) and variant QC (allele fraction, mapping quality, mismatches)		Variant QC (AF filter and read depth threshold); specific VAF filtering (<0.3 or >0.7) to account for potential CHIP in *NF1* and *TP53*		Sample QC (kinship filtering; exclusion of sex-discrepant individuals) and variant QC (read depth and callability fraction threshold); specific VAF filtering (<0.3 or >0.7) to account for potential CHIP in *NF1* and *TP53*	
Criteria for pathogenic variant (PV) inclusion		PTV tier only in primary reported analysis. Missense tier of ClinVar 2* P/LP variants additionally included in this analysis		Combined PTV and pathogenic non-PTV tier (based on classifications from ClinVar, diagnostic laboratories and ENIGMA)		PTV tier + pathogenic non-PTV tier (ClinVar 2* P/LP)	
Logistic regression covariates		Country, ethnicity group (Asian studies) or geographical stratum (UK), depending on study		Study, age, first-degree family history and race/ethnic group		Top-level ethnicity and age at enrolment	

Supplementary Table 2a: Full data for 37 putative BCSGs from each study. OR = odds ratio, PTV = protein-truncating variant, PV = pathogenic variant. For each gene, the odds ratio, upper and lower 95% confidence intervals of the odds ratio, standard error, and *p*-value derived from the Wald test statistic were calculated.

Constituent datasets

Gene	BRIDGES (PTVs only, as published)									BRIDGES (pathogenic missense only)								
	Cases w/ variant	Controls w/ variant	Cases w/o variant	Controls w/o variant	OR	Lower 95% CI	Upper 95% CI	Standard error	*p*-value	Cases w/ variant	Controls w/ variant	Cases w/o variant	Controls w/o variant	OR	Lower 95% CI	Upper 95% CI	Standard error	*p*-value
ABRAXAS1	17	19	48809	50684	0.98	0.50	1.94	0.35	0.96	0	0	48826	50703	NA	NA	NA	NA	NA
AKT1	3	6	48823	50697	0.47	0.12	1.93	0.71	0.29	0	0	48826	50703	NA	NA	NA	NA	NA
ATM	294	150	48532	50553	2.10	1.71	2.57	0.10	9.20E-13	56	25	48770	50678	2.39	1.49	3.84	0.24	4.46E-04
BABAM2	7	9	48819	50694	0.62	0.23	1.71	0.51	0.36	0	0	48826	50703	NA	NA	NA	NA	NA
BARD1	62	32	48764	50671	2.09	1.35	3.23	0.22	9.80E-04	0	0	48826	50703	NA	NA	NA	NA	NA
BLM	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA
BRCA1	515	58	48311	50645	10.57	8.02	13.93	0.14	1.10E-62	97	19	48729	50684	6.03	3.69	9.86	0.25	2.87E-11
BRCA2	754	135	48072	50568	5.85	4.85	7.06	0.10	2.20E-75	54	15	48772	50688	3.73	2.10	6.60	0.29	6.19E-06
BRIP1	86	75	48740	50628	1.11	0.80	1.53	0.17	0.54	3	1	48823	50702	2.90	0.30	27.91	1.15	0.33
CDH1	11	12	48815	50691	0.86	0.37	1.98	0.43	0.72	1	2	48825	50701	0.47	0.04	5.17	1.22	0.59
CDKN2A	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA
CHEK2 (all variants)	704	315	48122	50388	2.54	2.21	2.91	0.07	3.10E-39	57	31	48769	50672	1.76	1.13	2.72	0.22	3.74E-03
CHEK2 (c.1100delC only)	548	245	48278	50458	2.66	2.27	3.11	0.08	1.10E-33	0	0	48826	50703	NA	NA	NA	NA	NA
CHEK2 (excluding c.1100delC)	156	70	48670	50633	2.13	1.60	2.84	0.15	3.00E-07	57	31	48769	50672	1.76	1.13	2.72	0.22	3.74E-03
EPCAM	14	19	48812	50684	0.73	0.36	1.49	0.36	0.39	0	0	48826	50703	NA	NA	NA	NA	NA
ERCC3	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA
FANCC	71	65	48755	50638	1.26	0.89	1.79	0.18	0.2	3	0	48823	50703	8.07	0.42	156.32	1.51	0.19
FANCM	302	300	48524	50403	1.06	0.90	1.26	0.09	0.48	0	0	48826	50703	NA	NA	NA	NA	NA
GEN1	31	43	48795	50660	0.66	0.41	1.06	0.24	0.088	0	0	48826	50703	NA	NA	NA	NA	NA
MEN1	2	5	48824	50698	0.37	0.07	1.97	0.85	0.24	1	1	48825	50702	0.93	0.06	14.79	1.41	0.98
MLH1	5	9	48821	50694	0.58	0.19	1.77	0.57	0.34	3	4	48823	50699	0.78	0.18	3.50	0.76	0.74
MRE11	48	55	48778	50648	0.88	0.59	1.32	0.21	0.54	0	0	48826	50703	NA	NA	NA	NA	NA
MSH2	13	13	48813	50690	1.06	0.47	2.36	0.41	0.89	1	4	48825	50699	0.26	0.03	2.37	1.12	0.23
MSH6	39	23	48787	50680	1.96	1.15	3.33	0.27	0.013	7	7	48819	50696	1.16	0.41	3.29	0.53	0.94
NBN	90	103	48736	50600	0.90	0.67	1.20	0.15	0.48	0	0	48826	50703	NA	NA	NA	NA	NA
NF1	31	17	48826	50703	1.76	0.96	3.21	0.31	0.068	6	4	48826	50703	1.45	0.41	5.13	0.65	0.49
PALB2	274	55	48552	50648	5.02	3.73	6.76	0.15	1.60E-26	0	0	48826	50703	NA	NA	NA	NA	NA
PIK3CA	3	12	48823	50691	0.21	0.06	0.75	0.64	0.016	0	1	48826	50702	0.28	0.01	6.87	1.63	0.52
PMS2	40	36	48786	50667	1.16	0.73	1.85	0.24	0.53	19	24	48807	50679	0.83	0.45	1.51	0.31	0.52
PTEN	14	6	48812	50697	2.25	0.85	6.00	0.50	0.1	2	1	48824	50702	1.93	0.17	21.27	1.22	0.55
RAD50	120	121	48706	50582	1.08	0.83	1.40	0.13	0.57	0	0	48826	50703	NA	NA	NA	NA	NA
RAD51C	54	26	48772	50677	1.93	1.20	3.11	0.24	7.00E-03	2	0	48824	50703	4.64	0.22	96.73	1.55	0.29
RAD51D	51	25	48775	50678	1.80	1.11	2.93	0.25	0.018	0	0	48826	50703	NA	NA	NA	NA	NA
RECQL	103	120	48723	50583	0.84	0.64	1.10	0.14	0.21	0	0	48826	50703	NA	NA	NA	NA	NA
RINT1	32	49	48794	50654	0.72	0.46	1.14	0.23	0.17	0	0	48826	50703	NA	NA	NA	NA	NA
SLX4	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA
STK11	6	5	48820	50698	1.60	0.48	5.28	0.61	0.44	0	0	48826	50703	NA	NA	NA	NA	NA
TP53	7	2	48826	50703	3.06	0.63	14.91	0.81	0.17	27	7	48826	50703	3.37	1.47	7.75	0.42	1.07E-03
XRCC2	15	18	48811	50685	0.96	0.47	1.93	0.36	0.9	0	0	48826	50703	NA	NA	NA	NA	NA

Gene	CARRIERS (all PVs)									UK Biobank (PTVs only)								
	Cases w/ variant	Controls w/ variant	Cases w/o variant	Controls w/o variant	OR	Lower 95% CI	Upper 95% CI	Standard error	*p*-value	Cases w/ variant	Controls w/ variant	Cases w/o variant	Controls w/o variant	OR	Lower 95% CI	Upper 95% CI	Standard error	*p*-value
ABRAXAS1	NA	NA	NA	NA	NA	NA	NA	NA	NA	10	93	20314	229604	1.17	0.61	2.26	0.33	0.63
AKT1	NA	NA	NA	NA	NA	NA	NA	NA	NA	0	13	20324	229684	2.72E-04	2.18E-32	3.40E+24	33.01	0.80
ATM	253	134	31994	32410	1.82	1.46	2.27	0.11	1.04E-07	118	602	20206	229095	2.24	1.83	2.73	0.10	2.22E-15
BABAM2	NA	NA	NA	NA	NA	NA	NA	NA	NA	6	62	20318	229635	1.13	0.49	2.62	0.43	0.78
BARD1	49	35	32198	32509	1.37	0.87	2.16	0.23	0.18	39	125	20285	229572	3.50	2.43	5.03	0.19	1.36E-11
BLM	104	87	32143	32457	1.19	0.89	1.59	0.15	0.24	59	526	20265	229171	1.25	0.95	1.64	0.14	0.10
BRCA1	275	37	31972	32507	7.62	5.33	11.27	0.19	2.13E-26	94	131	20230	229566	9.22	7.04	12.06	0.14	5.41E-59
BRCA2	417	78	31830	32466	5.23	4.09	6.77	0.13	6.75E-38	254	525	20070	229172	5.79	4.98	6.75	0.08	1.83E-113
BRIP1	69	52	32178	32492	1.35	0.93	1.98	0.19	0.12	46	409	20278	229288	1.28	0.94	1.74	0.16	0.11
CDH1	17	6	32230	32538	2.50	1.01	7.07	0.50	0.06	7	23	20317	229674	3.83	1.63	9.00	0.44	2.06E-03
CDKN2A	8	5	32239	32539	1.51	0.47	5.22	0.61	0.5	1	11	20323	229686	1.23	0.16	9.61	1.05	0.85
CHEK2 (all variants)	349	138	31898	32406	2.47	2.02	3.05	0.11	7.81E-18	252	1219	20072	228478	2.35	2.04	2.69	0.07	4.50E-34
CHEK2 (c.1100delC only)	NA	NA	NA	NA	NA	NA	NA	NA	NA	189	878	20135	228819	2.43	2.07	2.85	0.08	5.87E-28
CHEK2 (excluding c.1100delC)	NA	NA	NA	NA	NA	NA	NA	NA	NA	63	343	20261	229354	2.09	1.60	2.74	0.14	9.16E-08
EPCAM	NA	NA	NA	NA	NA	NA	NA	NA	NA	4	51	20320	229646	0.81	0.29	2.24	0.52	0.68
ERCC3	56	83	32191	32461	0.71	0.50	1.01	0.18	0.06	42	523	20282	229174	0.92	0.67	1.25	0.16	0.58
FANCC	75	104	32172	32440	0.75	0.55	1.01	0.16	0.06	32	341	20292	229356	1.05	0.73	1.51	0.19	0.80
FANCM	51	46	32196	32498	1.14	0.76	1.74	0.21	0.52	118	1063	20206	228634	1.26	1.04	1.52	0.10	0.02
GEN1	NA	NA	NA	NA	NA	NA	NA	NA	NA	21	278	20303	229419	0.86	0.55	1.34	0.23	0.49
MEN1	NA	NA	NA	NA	NA	NA	NA	NA	NA	1	4	20323	229693	3.25	0.36	29.58	1.13	0.29
MLH1	10	3	32237	32541	3.36	0.93	12.23	0.66	0.07	6	79	20318	229618	0.90	0.39	2.08	0.43	0.81
MRE11	25	32	32222	32512	0.69	0.38	1.20	0.29	0.19	17	237	20307	229460	0.79	0.48	1.30	0.25	0.35
MSH2	7	5	32240	32539	1.28	0.38	4.47	0.63	0.68	7	80	20317	229617	0.97	0.45	2.11	0.40	0.95
MSH6	39	32	32208	32512	1.13	0.70	1.83	0.25	0.63	20	193	20304	229504	1.18	0.74	1.87	0.24	0.48
NBN	57	51	32190	32493	1.05	0.71	1.56	0.20	0.81	31	407	20293	229290	0.85	0.59	1.23	0.19	0.38
NF1	19	11	32247	32544	1.93	0.91	4.31	0.40	0.09	9	57	20315	229640	2.15	1.16	4.00	0.32	0.02
PALB2	148	38	32099	32506	3.83	2.68	5.63	0.19	1.33E-12	110	314	20214	229383	4.05	3.25	5.05	0.11	9.77E-36
PIK3CA	NA	NA	NA	NA	NA	NA	NA	NA	NA	2	8	20322	229689	3.18	0.67	15.17	0.80	0.15
PMS2	NA	NA	NA	NA	NA	NA	NA	NA	NA	48	520	20276	229177	1.16	0.86	1.57	0.15	0.32
PTEN	8	3	32239	32541	2.69	0.71	10.15	0.68	NA	1	4	20323	229693	3.51	0.39	31.71	1.12	0.26
RAD50	57	82	32190	32462	0.73	0.51	1.04	0.18	0.08	45	556	20279	229141	0.93	0.69	1.27	0.16	0.66
RAD51C	41	35	32206	32509	1.20	0.75	1.93	0.24	0.44	9	91	20315	229606	1.19	0.60	2.36	0.35	0.63
RAD51D	26	14	32221	32530	1.72	0.88	3.51	0.35	0.12	17	110	20307	229587	1.74	1.04	2.91	0.26	0.03
RECQL	74	69	32173	32475	1.03	0.74	1.45	0.17	0.86	50	492	20274	229205	1.18	0.88	1.59	0.15	0.26
RINT1	24	28	32223	32516	0.80	0.45	1.41	0.29	0.44	8	117	20316	229580	0.76	0.37	1.56	0.37	0.46
SLX4	44	41	32203	32503	1.03	0.66	1.60	0.23	0.91	24	321	20300	229376	0.85	0.56	1.29	0.21	0.45
STK11	NA	NA	NA	NA	NA	NA	NA	NA	NA	0	1	20324	229696	2.54E-03	9.77E-41	6.58E+34	43.95	0.89
TP53	19	2	32247	32544	9.59	2.23	41.19	0.74	2.36E-03	1	7	20323	229690	1.59	0.19	12.99	1.07	0.67
XRCC2	27	21	32220	32523	1.19	0.67	2.17	0.30	0.55	2	16	20322	229681	1.43	0.33	6.25	0.75	0.64

Gene	UK Biobank (non-PTV-PVs only)									UK Biobank (PTVs + non-PTV-PVs)								
	Cases w/ variant	Controls w/ variant	Cases w/o variant	Controls w/o variant	OR	Lower 95% CI	Upper 95% CI	Standard error	*p*-value	Cases w/ variant	Controls w/ variant	Cases w/o variant	Controls w/o variant	OR	Lower 95% CI	Upper 95% CI	Standard error	*p*-value
ABRAXAS1	0	0	20324	229697	NA	NA	NA	NA	NA	10	93	20314	229604	1.17	0.61	2.26	0.33	0.63
AKT1	0	0	20324	229697	NA	NA	NA	NA	NA	0	13	20324	229684	2.72E-04	2.18E-32	3.40E+24	33.01	0.80
ATM	46	179	20278	229518	2.93	2.11	4.06	0.17	1.13E-10	164	781	20160	228916	2.40	2.03	2.85	0.09	5.91E-24
BABAM2	0	0	20324	229697	NA	NA	NA	NA	NA	6	62	20318	229635	1.13	0.49	2.62	0.43	0.78
BARD1	1	1	20323	229696	8.01	0.50	128.63	1.42	0.14	40	126	20284	229571	3.54	2.47	5.07	0.18	4.99E-12
BLM	1	10	20323	229687	1.09	0.14	8.54	1.05	9.36E-01	60	536	20264	229161	1.25	0.95	1.63	0.14	0.11
BRCA1	18	31	20306	229666	6.71	3.74	12.04	0.30	1.85E-10	112	162	20212	229535	8.72	6.83	11.13	0.12	1.17E-67
BRCA2	20	30	20304	229667	7.71	4.35	13.67	0.29	2.79E-12	274	555	20050	229142	5.91	5.10	6.84	0.07	5.24E-124
BRIP1	3	49	20321	229648	0.70	0.22	2.24	0.60	0.54	49	458	20275	229239	1.22	0.91	1.64	0.15	0.19
CDH1	2	4	20322	229693	5.57	1.00	31.06	0.88	0.05	9	27	20315	229670	4.11	1.92	8.81	0.39	2.79E-04
CDKN2A	19	103	20305	229594	2.05	1.25	3.36	0.25	4.27E-03	20	114	20304	229583	1.98	1.23	3.20	0.24	5.05E-03
CHEK2 (all variants)	23	119	20301	229578	2.12	1.35	3.31	0.23	1.08E-03	275	1338	20049	228359	2.33	2.04	2.65	0.07	2.19E-36
CHEK2 (c.1100delC only)	0	0	20324	229697	NA	NA	NA	NA	NA	189	878	20135	228819	2.43	2.07	2.85	0.08	5.87E-28
CHEK2 (excluding c.1100delC)	23	119	20301	229578	2.12	1.35	3.31	0.23	1.08E-03	86	462	20238	229235	2.10	1.66	2.65	0.12	3.72E-10
EPCAM	0	27	20324	229670	1.03E-04	8.66E-37	1.23E+28	37.68	0.81	4	78	20320	229619	0.54	0.20	1.47	0.51	0.22
ERCC3	0	0	20324	229697	NA	NA	NA	NA	NA	42	523	20282	229174	0.92	0.67	1.25	0.16	0.58
FANCC	8	95	20316	229602	0.90	0.44	1.85	0.37	0.77	40	436	20284	229261	1.01	0.73	1.40	0.17	0.93
FANCM	0	0	20324	229697	NA	NA	NA	NA	NA	118	1063	20206	228634	1.26	1.04	1.52	0.10	0.02
GEN1	0	0	20324	229697	NA	NA	NA	NA	NA	21	278	20303	229419	0.86	0.55	1.34	0.23	0.49
MEN1	3	2	20321	229695	14.08	2.34	84.72	0.92	3.87E-03	4	6	20320	229691	7.39	2.06	26.53	0.65	2.17E-03
MLH1	2	23	20322	229674	0.95	0.22	4.04	0.74	0.94	8	102	20316	229595	0.92	0.44	1.88	0.37	0.81
MRE11	0	0	20324	229697	NA	NA	NA	NA	NA	17	237	20307	229460	0.79	0.48	1.30	0.25	0.35
MSH2	0	7	20324	229690	2.88E-04	2.45E-42	3.37E+34	44.72	0.86	7	87	20317	229610	0.89	0.41	1.94	0.39	0.78
MSH6	3	50	20321	229647	0.66	0.21	2.13	0.60	0.49	23	243	20301	229454	1.07	0.70	1.64	0.22	0.75
NBN	1	6	20323	229691	1.81	0.21	15.31	1.09	0.59	32	413	20292	229284	0.86	0.60	1.24	0.18	0.43
NF1	4	13	20320	229684	3.38	1.23	9.27	0.51	1.78E-02	13	70	20311	229627	2.41	1.42	4.07	0.27	1.08E-03
PALB2	11	24	20313	229673	5.20	2.53	10.70	0.37	7.59E-06	121	338	20203	229359	4.14	3.35	5.10	0.11	5.08E-40
PIK3CA	0	0	20324	229697	NA	NA	NA	NA	NA	2	8	20322	229689	3.18	0.67	15.17	0.80	0.15
PMS2	27	230	20297	229467	1.38	0.92	2.06	0.20	0.12	75	746	20249	228951	1.24	0.98	1.57	0.12	0.08
PTEN	2	7	20322	229690	3.78	0.77	18.48	0.81	0.10	3	11	20321	229686	3.69	1.02	13.34	0.66	0.05
RAD50	0	0	20324	229697	NA	NA	NA	NA	NA	45	556	20279	229141	0.93	0.69	1.27	0.16	0.66
RAD51C	5	20	20319	229677	2.98	1.11	8.02	0.50	0.03	14	111	20310	229586	1.51	0.86	2.64	0.29	0.15
RAD51D	0	0	20324	229697	NA	NA	NA	NA	NA	17	110	20307	229587	1.74	1.04	2.91	0.26	0.03
RECQL	0	0	20324	229697	NA	NA	NA	NA	NA	50	492	20274	229205	1.18	0.88	1.59	0.15	0.26
RINT1	0	0	20324	229697	NA	NA	NA	NA	NA	8	117	20316	229580	0.76	0.37	1.56	0.37	0.46
SLX4	0	0	20324	229697	NA	NA	NA	NA	NA	24	321	20300	229376	0.85	0.56	1.29	0.21	0.45
STK11	0	0	20324	229697	NA	NA	NA	NA	NA	0	1	20324	229696	2.54E-03	9.77E-41	6.58E+34	43.95	0.89
TP53	1	5	20323	229692	2.31	0.27	19.85	1.10	0.45	2	12	20322	229685	1.88	0.42	8.45	0.77	0.41
XRCC2	0	0	20324	229697	NA	NA	NA	NA	NA	2	16	20322	229681	1.43	0.33	6.25	0.75	0.64

Supplementary Table 2b: Full data for 37 putative BCSGs from meta-analyses. PTV = protein-truncating variant, PV = pathogenic variant. For each gene, the odds ratio (OR), upper and lower 95% confidence intervals (CIs) for the odds ratio, standard error, and *p*-value based on combined standard error were calculated. Heterogeneity estimates are presented where genes were analysed in all three studies (as unreliable where n=2), and comprise Cochran's Q (a chi-squared distributed measure of heterogeneity) and I^2 (a measure of heterogeneity than Cochran's Q more robust at low n, evaluating the percentage of variability in 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)

	Meta-Analyses																									
	All available datasets (PTVs + pathogenic missense/non-PTV-PVs)							BRIDGES (PTVs + pathogenic missense)					BRIDGES (PTVs only) + CARRIERS (all PVs)					BRIDGES (PTVs only) + UK Biobank (PTVs only)					BRIDGES (pathogenic missense only) + UK Biobank (non-PTV-PVs only)				
						Heterogeneity																					
Gene	OR	Lower 95% CI	Upper 95% CI	Standard error	*p*-value	Cochran's Q	I^2 (%)	OR	Lower 95% CI	Upper 95% CI	Standard error	*p*-value	OR	Lower 95% CI	Upper 95% CI	Standard error	*p*-value	OR	Lower 95% CI	Upper 95% CI	Standard error	*p*-value	OR	Lower 95% CI	Upper 95% CI	Standard error	*p*-value
ABRAXAS1	1.08	0.67	1.72	0.24	0.76	NA	NA	0.98	0.50	1.94	0.35	0.96	0.98	0.50	1.94	0.35	0.96	1.08	0.67	1.72	0.24	0.76	NA	NA	NA	NA	NA
AKT1	0.47	0.12	1.88	0.71	0.28	NA	NA	0.47	0.12	1.93	0.71	0.29	0.47	0.12	1.93	0.71	0.29	0.47	0.12	1.88	0.71	0.28	NA	NA	NA	NA	NA
ATM	2.16	1.93	2.41	0.06	2.83E-43	3.81	47.55	2.14	1.78	2.58	0.10	1.33E-15	1.97	1.69	2.28	0.08	8.51E-19	2.17	1.88	2.50	0.07	1.57E-26	2.74	2.10	3.59	0.14	1.75E-13
BABAM2	0.88	0.46	1.68	0.33	0.70	NA	NA	0.62	0.23	1.71	0.51	0.36	0.62	0.23	1.71	0.51	0.36	0.88	0.46	1.68	0.33	0.70	NA	NA	NA	NA	NA
BARD1	2.34	1.85	2.97	0.12	1.71E-12	10.71	81.33	2.09	1.35	3.23	0.22	9.80E-04	1.71	1.25	2.34	0.16	8.69E-04	2.83	2.14	3.74	0.14	2.53E-13	8.01	0.78	82.33	1.42	0.14
BLM	1.22	1.00	1.49	0.10	0.047	NA	NA	NA	NA	NA	NA	NA	1.19	0.89	1.59	0.15	0.24	1.25	1.00	1.57	0.14	0.10	1.09	0.19	6.14	1.05	0.94
BRCA1	8.73	7.47	10.20	0.08	2.07E-163	0.72	0	9.24	7.26	11.76	0.12	2.97E-73	9.42	7.54	11.76	0.11	4.06E-87	9.85	8.13	11.94	0.10	6.54E-120	6.30	4.32	9.18	0.19	9.60E-22
BRCA2	5.68	5.13	6.30	0.05	9.48E-238	0.71	0	5.60	4.68	6.69	0.09	6.75E-80	5.62	4.84	6.53	0.08	6.90E-112	5.82	5.17	6.55	0.06	1.99E-187	5.36	3.57	8.03	0.21	4.45E-16
BRIP1	1.22	1.01	1.47	0.10	0.04	0.49	0	1.13	0.82	1.56	0.16	0.45	1.21	0.94	1.54	0.13	0.14	1.20	0.96	1.50	0.11	0.11	0.94	0.33	2.65	0.53	0.91
CDH1	2.01	1.25	3.24	0.24	4.20E-03	8.70	77.02	0.81	0.36	1.78	0.40	0.59	1.36	0.72	2.56	0.32	0.35	1.79	0.98	3.26	0.31	0.06	2.41	0.60	9.75	0.71	0.22
CDKN2A	1.91	1.22	2.98	0.23	4.32E-03	NA	NA	NA	NA	NA	NA	NA	1.51	0.47	5.22	0.61	0.50	1.23	0.22	6.90	1.05	0.85	2.05	1.36	3.10	0.25	4.27E-03
CHEK2 (all variants)	2.40	2.21	2.62	0.04	1.06E-91	0.41	0	2.46	2.15	2.80	0.07	4.36E-41	2.52	2.25	2.82	0.06	2.23E-56	2.44	2.21	2.69	0.05	2.14E-72	1.92	1.40	2.63	0.16	4.40E-05
CHEK2 (c.1100delC only)	2.54	2.27	2.85	0.06	3.26E-60	NA	NA	2.66	2.27	3.11	0.08	1.10E-33	2.66	2.27	3.11	0.08	1.10E-33	2.54	2.27	2.85	0.06	3.26E-60	NA	NA	NA	NA	NA
CHEK2 (excluding c.1100delC)	2.06	1.74	2.43	0.09	2.51E-17	NA	NA	2.01	1.58	2.55	0.12	1.19E-08	2.13	1.60	2.84	0.15	3.00E-07	2.11	1.73	2.57	0.10	1.08E-13	1.92	1.40	2.63	0.16	4.40E-05
EPCAM	0.66	0.37	1.18	0.30	0.16	NA	NA	0.73	0.36	1.49	0.36	0.39	0.73	0.36	1.49	0.36	0.39	0.75	0.42	1.35	0.30	0.34	1.03E-04	1.24E-31	8.62E+22	37.68	0.81
ERCC3	0.82	0.65	1.03	0.12	0.09	NA	NA	NA	NA	NA	NA	NA	0.71	0.50	1.01	0.18	0.06	0.92	0.70	1.19	0.16	0.58	NA	NA	NA	NA	NA
FANCC	0.97	0.81	1.17	0.10	0.75	5.45	63.32	1.29	0.91	1.83	0.18	0.15	0.94	0.75	1.18	0.12	0.58	1.15	0.90	1.48	0.13	0.27	1.02	0.50	2.05	0.36	0.97
FANCM	1.14	1.01	1.29	0.06	0.03	1.69	0	1.06	0.90	1.26	0.09	0.48	1.07	0.92	1.25	0.08	0.39	1.14	1.01	1.29	0.06	0.04	NA	NA	NA	NA	NA
GEN1	0.76	0.55	1.05	0.17	0.09	NA	NA	0.66	0.41	1.06	0.24	0.09	0.66	0.41	1.06	0.24	0.09	0.76	0.55	1.05	0.17	0.09	NA	NA	NA	NA	NA
MEN1	2.18	0.84	5.64	0.49	0.11	NA	NA	0.47	0.11	1.97	0.73	0.30	0.37	0.07	1.97	0.85	0.24	0.82	0.22	3.09	0.68	0.76	6.30	1.40	28.42	0.77	0.02
MLH1	1.00	0.60	1.68	0.26	0.99	4.37	54.27	0.65	0.26	1.58	0.46	0.34	1.23	0.53	2.86	0.43	0.63	0.77	0.40	1.50	0.34	0.45	0.86	0.30	2.45	0.53	0.78
MRE11	0.81	0.61	1.06	0.14	0.12	0.47	0	0.88	0.59	1.32	0.21	0.54	0.81	0.58	1.13	0.17	0.22	0.84	0.62	1.15	0.16	0.28	NA	NA	NA	NA	NA
MSH2	0.95	0.58	1.56	0.25	0.84	0.27	0	0.90	0.42	1.92	0.39	0.78	1.12	0.57	2.20	0.34	0.74	1.01	0.58	1.77	0.29	0.96	0.26	0.03	2.36	1.12	0.23
MSH6	1.27	0.98	1.66	0.14	0.08	2.65	24.46	1.76	1.09	2.83	0.24	0.02	1.45	1.01	2.07	0.18	0.04	1.47	1.04	2.08	0.18	0.03	0.90	0.41	1.97	0.40	0.79
NBN	0.92	0.76	1.12	0.10	0.43	0.57	0	0.90	0.67	1.20	0.15	0.48	0.95	0.75	1.20	0.12	0.67	0.88	0.70	1.11	0.12	0.27	1.81	0.30	10.86	1.09	0.59
NF1	2.01	1.43	2.83	0.17	5.62E-05	0.83	0	1.70	0.98	2.93	0.28	0.06	1.82	1.13	2.94	0.24	0.01	1.94	1.26	2.99	0.22	2.69E-03	2.43	1.11	5.35	0.40	0.03
PALB2	4.30	3.68	5.03	0.08	2.59E-75	1.55	0.00	5.02	3.73	6.76	0.15	1.60E-26	4.52	3.58	5.70	0.12	3.77E-37	4.37	3.66	5.21	0.09	3.87E-60	5.20	2.84	9.53	0.37	7.59E-06
PIK3CA	0.57	0.22	1.47	0.48	0.25	NA	NA	0.22	0.07	0.71	0.60	0.01	0.21	0.06	0.75	0.64	0.02	0.61	0.23	1.64	0.50	0.33	0.28	0.02	4.11	1.63	0.52
PMS2	1.17	0.96	1.43	0.10	0.13	NA	NA	1.02	0.71	1.48	0.19	0.91	1.16	0.73	1.85	0.24	0.53	1.16	0.90	1.49	0.13	0.24	1.18	0.84	1.64	0.17	0.34
PTEN	2.63	1.38	5.02	0.33	3.38E-03	0.41	0.00	2.20	0.89	5.44	0.46	0.09	2.40	1.09	5.26	0.40	0.03	2.42	0.99	5.91	0.46	0.05	3.08	0.82	11.58	0.68	0.10
RAD50	0.94	0.79	1.12	0.09	0.48	3.02	33.77	1.08	0.83	1.40	0.13	0.57	0.94	0.76	1.16	0.11	0.58	1.02	0.83	1.24	0.10	0.88	NA	NA	NA	NA	NA
RAD51C	1.53	1.15	2.04	0.15	3.48E-03	2.13	6	1.97	1.23	3.16	0.24	4.69E-03	1.52	1.09	2.12	0.17	0.01	1.65	1.11	2.44	0.20	0.01	3.11	1.22	7.97	0.48	0.02
RAD51D	1.76	1.29	2.41	0.16	4.18E-04	0.01	0.00	1.80	1.11	2.93	0.25	0.02	1.77	1.19	2.64	0.20	4.72E-03	1.77	1.24	2.52	0.18	1.49E-03	NA	NA	NA	NA	NA
RECQL	1.00	0.84	1.18	0.09	0.96	2.91	31	0.84	0.64	1.10	0.14	0.21	0.91	0.74	1.12	0.11	0.38	0.98	0.81	1.20	0.10	0.88	NA	NA	NA	NA	NA
RINT1	0.75	0.55	1.03	0.16	0.08	0.08	0.00	0.72	0.46	1.14	0.23	0.17	0.75	0.53	1.07	0.18	0.11	0.73	0.50	1.07	0.20	0.11	NA	NA	NA	NA	NA
SLX4	0.93	0.69	1.26	0.15	0.65	NA	NA	NA	NA	NA	NA	NA	1.03	0.66	1.60	0.23	0.91	0.85	0.60	1.21	0.21	0.45	NA	NA	NA	NA	NA
STK11	1.60	0.48	5.30	0.61	0.44	NA	NA	1.60	0.48	5.28	0.61	0.44	1.60	0.48	5.28	0.61	0.44	1.60	0.48	5.30	0.61	0.44	NA	NA	NA	NA	NA
TP53	3.62	1.98	6.61	0.31	2.79E-05	2.50	20	3.30	1.58	6.89	0.38	1.46E-03	5.68	1.94	16.58	0.55	1.49E-03	2.41	0.68	8.54	0.64	0.17	3.21	1.48	6.97	0.40	3.19E-03
XRCC2	1.12	0.72	1.72	0.22	0.62	0.33	0.00	0.96	0.47	1.93	0.36	0.90	1.09	0.69	1.71	0.23	0.71	1.03	0.55	1.95	0.33	0.92	NA	NA	NA	NA	NA

Supplementary Table 3: Pathogenic variant (PV) frequencies in all 37 surveyed BC-MGPT and putative BC genes across cases and controls in constituent studies and the combined meta-analysis. 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 in 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*-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. BCSG = breast cancer susceptibility gene; PTV = protein-truncating variant; PV = pathogenic variant; UKB = UK Biobank.

		Cases					Controls				
	Gene	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%)	2.49E-18	**999 (0.99%; 101)**	77 (0.15%)	37 (0.11%)	162 (0.070%)	4.43E-08	**276 (0.09%; 1134)**
	BRCA2	808 (1.65%)	417 (1.29%)	274 (1.35%)	3.71E-05	**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	***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%)	1.61E-03	**193 (0.06%; 1621)**
	***CHEK2* (all variants)**	761 (1.56%)	349 (1.08%)	275 (1.35%)	7.79E-08	**1385 (1.37%; 73)**	346 (0.68%)	138 (0.42%)	1338 (0.58%)	1.09E-05	**1822 (0.58%; 172)**
	***CHEK2* (c.1100delC only)**	548 (1.12%)	NA	189 (0.93%)	0.03	**737 (1.07%; 94)**	245 (0.48%)	NA	878 (0.38%)	1.29E-03	**1123 (0.4%; 250)**
	***CHEK2* (excluding c.1100delC)**	213 (0.44%)	NA	86 (0.42%)	0.86	**299 (0.43%; 231)**	101 (0.20%)	NA	462 (0.20%)	0.97	**563 (0.2%; 498)**
	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%)	1.04E-04	**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	***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)**
	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)**
	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%; 14902)**
	STK11	6 (0.012%)	NA	0 (0%)	0.19	**6 (0.01%; 11525)**	5 (0.0099%)	NA	1 (0.00044%)	9.85E-04	**6 (0%; 46733)**
	TP53	34 (0.070%)	19 (0.059%)	2 (0.0098%)	8.01E-03	**55 (0.05%; 1844)**	9 (0.018%)	2 (0.0061%)	12 (0.0052%)	0.01	**23 (0.01%; 13606)**
	EPCAM	14 (0.029%)	NA	4 (0.020%)	0.61	**18 (0.03%; 3842)**	19 (0.037%)	NA	78 (0.034%)	0.80	**97 (0.03%; 2891)**

Mis-match repair (MMR) genes	***MLH1***	8 (0.016%)	10 (0.031%)	8 (0.039%)	0.17	**26 (0.03%; 3900)**	13 (0.026%)	3 (0.0092%)	102 (0.044%)	2.88E-03	**118 (0.04%; 2652)**
	MSH2	14 (0.029%)	7 (0.022%)	7 (0.034%)	0.68	**28 (0.03%; 3621)**	17 (0.034%)	5 (0.015%)	87 (0.038%)	0.12	**109 (0.03%; 2871)**
	MSH6	46 (0.094%)	39 (0.12%)	23 (0.11%)	0.49	**108 (0.11%; 939)**	30 (0.059%)	32 (0.098%)	243 (0.10%)	9.68E-03	**305 (0.1%; 1026)**
	PMS2	59 (0.12%)	NA	75 (0.37%)	2.64E-11	**134 (0.19%; 516)**	60 (0.12%)	NA	746 (0.32%)	5.59E-15	**806 (0.29%; 348)**
Other putative BCSGs	***ABRAXAS1***	17 (0.035%)	NA	10 (0.049%)	0.51	**27 (0.04%; 2561)**	19 (0.037%)	NA	93 (0.040%)	0.85	**112 (0.04%; 2504)**
	AKT1	3 (0.0061%)	NA	0 (0%)	0.56	**3 (0%; 23050)**	6 (0.012%)	NA	13 (0.0057%)	0.22	**19 (0.01%; 14758)**
	BABAM2	7 (0.014%)	NA	6 (0.030%)	0.31	**13 (0.02%; 5319)**	9 (0.018%)	NA	62 (0.027%)	0.30	**71 (0.03%; 3949)**
	BLM	NA	104 (0.32%)	60 (0.30%)	0.64	**164 (0.31%; 321)**	NA	87 (0.27%)	536 (0.23%)	0.26	**623 (0.24%; 421)**
	BRIP1	89 (0.18%)	69 (0.21%)	49 (0.24%)	0.26	**207 (0.2%; 490)**	76 (0.15%)	52 (0.16%)	458 (0.20%)	0.03	**586 (0.19%; 534)**
	CDKN2A	NA	8 (0.025%)	20 (0.098%)	7.58E-04	**28 (0.05%; 1878)**	NA	5 (0.015%)	114 (0.050%)	9.95E-03	**119 (0.05%; 2204)**
	ERCC3	NA	56 (0.17%)	42 (0.20%)	0.45	**98 (0.19%; 536)**	NA	83 (0.26%)	523 (0.23%)	0.37	**606 (0.23%; 433)**
	FANCC	74 (0.15%)	75 (0.23%)	40 (0.20%)	0.03	**189 (0.19%; 536)**	65 (0.13%)	104 (0.32%)	436 (0.19%)	5.14E-09	**605 (0.19%; 517)**
	FANCM	302 (0.62%)	51 (0.16%)	118 (0.58%)	1.15E-21	**471 (0.46%; 215)**	300 (0.59%)	46 (0.14%)	1063 (0.46%)	7.36E-21	**1409 (0.45%; 222)**
	GEN1	31 (0.063%)	NA	21 (0.10%)	0.11	**52 (0.08%; 1330)**	43 (0.085%)	NA	278 (0.12%)	0.03	**321 (0.11%; 874)**
	MEN1	3 (0.0061%)	NA	4 (0.020%)	0.21	**7 (0.01%; 9879)**	6 (0.012%)	NA	6 (0.0026%)	0.01	**12 (0%; 23367)**
	MRE11	48 (0.098%)	25 (0.078%)	17 (0.084%)	0.60	**90 (0.09%; 1127)**	55 (0.10%)	32 (0.098%)	237 (0.10%)	0.90	**324 (0.1%; 966)**
	NBN	90 (0.18%)	57 (0.18%)	32 (0.16%)	0.75	**179 (0.18%; 566)**	103 (0.20%)	51 (0.16%)	413 (0.18%)	0.29	**567 (0.18%; 552)**
	PIK3CA	3 (0.0061%)	NA	2 (0.0098%)	0.63	**5 (0.01%; 13830)**	13 (0.026%)	NA	8 (0.0035%)	8.04E-07	**21 (0.01%; 13352)**
	RAD50	120 (0.25%)	57 (0.18%)	45 (0.22%)	0.12	**222 (0.22%; 457)**	121 (0.24%)	82 (0.25%)	556 (0.24%)	0.93	**759 (0.24%; 412)**
	RECQL	103 (0.21%)	74 (0.23%)	50 (0.25%)	0.65	**227 (0.22%; 447)**	120 (0.24%)	69 (0.21%)	492 (0.21%)	0.60	**681 (0.22%; 460)**
	RINT1	32 (0.066%)	24 (0.074%)	8 (0.039%)	0.28	**64 (0.06%; 1584)**	49 (0.097%)	28 (0.086%)	117 (0.050%)	1.67E-04	**194 (0.06%; 1613)**
	SLX4	NA	44 (0.14%)	24 (0.12%)	0.66	**68 (0.13%; 773)**	NA	41 (0.13%)	321 (0.14%)	0.58	**362 (0.14%; 724)**
	XRCC2	15 (0.030%)	27 (0.084%)	2 (0.0098%)	6.85E-05	**44 (0.04%; 2304)**	18 (0.035%)	21 (0.065%)	16 (0.0070%)	8.42E-15	**55 (0.02%; 5690)**

Supplementary Table 4: ORs with 90% confidence intervals for 13 BC-MGPT genes in the combined meta-analysis. BC = breast cancer, PTV = protein-truncating variant. Where the lower 90% confidence interval is >4, values have been highlighted in dark grey; where the lower 90% confidence interval is >2, values have been highlighted in light grey, as per specification of high- and moderate-penetrance BC susceptibility genes by Easton et al. *The combined meta-analysis for *STK11* uses data from BRIDGES and UK Biobank only.

	101,397 BC cases 312,944 controls*
	OR (PTVs + non-PTVs)
BRCA1	8.73 (7.66-9.95)
BRCA2	5.68 (5.21-6.2)
PALB2	4.30 (3.78-4.90)
CHEK2	2.40 (2.24-2.58)
ATM	2.16 (1.97-2.37)
BARD1	2.34 (1.92-2.86)
RAD51C	1.53 (1.21-1.95)
RAD51D	1.76 (1.35-2.29)
TP53	3.62 (2.18-6.00)
PTEN	2.63 (1.53-4.52)
NF1	2.01 (1.51-2.68)
STK11	1.60 (0.58-4.37)
CDH1	2.01 (1.35-3.00)

Supplementary Table 5: Variant counts and odds ratios stratified by receptor status for each gene across BRIDGES, CARRIERS, and a combined meta-analysis of both datasets. BC = breast cancer, PTV = protein-truncating variant, TNT = triple-negative tumour. Odds ratios (ORs) have been provided with the lower and upper 95% confidence intervals. For the combined meta-analysis, data has been included from both BRIDGES (PTVs) and CARRIERS (PTVs and non-PTVs). Receptor status information was not available from UK Biobank.

	BRIDGES									CARRIERS									COMBINED								
	PTVs									PTVs + non-PTVs									PTVs + non-PTVs								
	All BC n = 48,826	ER-pos BC n = 30,466	ER-neg BC n = 7,766	TNT BC n = 2,841	Controls n = 50,703	OR (All)	OR (ER-pos)	OR (ER-neg)	OR (TNT)	All BC n = 32,247	ER-pos BC n = 18,428	ER-neg BC n = 3,805	TNT BC n = 1,463	Controls n = 32,544	OR (All)	OR (ER-pos)	OR (ER-neg)	OR (TNT)	All BC n = 81,073	ER-pos BC n = 48,894	ER-neg BC n = 11,571	TNT BC n = 4,304	Controls n = 83,247	OR (All)	OR (ER-pos)	OR (ER-neg)	OR (TNT)
BRCA1	515	120	269	165	58	10.57 (8.02-13.93)	3.92 (2.82-5.43)	35.32 (26.21-47.60)	56.80 (41.18-78.34)	275	73	114	65	37	7.62 (5.33-11.27)	3.39 (2.17-5.45)	26.33 (17.28-41.52)	42.88 (26.56-71.25)	790	193	383	230	95	9.42 (7.54-11.76)	3.73 (2.86-4.88)	32.18 (22.15-41.19)	52.23 (39.90-68.38)
BRCA2	754	446	149	74	135	5.85 (4.85-7.06)	5.69 (4.65-6.96)	7.53 (5.89-9.62)	11.19 (8.27-15.16)	417	201	82	30	78	5.23 (4.09-6.77)	4.66 (3.52-6.23)	8.89 (6.36-12.47)	9.7 (5.97-15.47)	1171	647	231	104	213	5.62 (4.84-6.53)	5.32 (4.52-6.28)	7.98 (6.54-9.73)	10.74 (8.32-13.87)
PALB2	274	152	56	29	55	5.02 (3.73-6.76)	4.45 (3.23-6.14)	6.72 (4.54-9.95)	10.36 (6.42-16.71)	148	64	42	24	38	3.83 (2.68-5.63)	3.13 (2.02-4.96)	9.22 (5.63-15.25)	13.03 (7.08-23.75)	422	216	98	53	93	4.52 (3.58-5.70)	3.95 (3.04-5.13)	7.58 (5.57-10.32)	11.31 (7.77-16.47)
CHEK2	704	481	67	16	315	2.54 (2.21-2.91)	2.67 (2.30-3.11)	1.64 (1.25-2.16)	1.06 (0.63-1.76)	349	205	20	8	138	2.47 (2.02-3.05)	2.60 (2.05-3.31)	1.4 (0.83-2.25)	1.63 (0.72-3.20)	1053	686	87	24	453	2.52 (2.25-2.82)	2.65 (2.33-3.01)	1.58 (1.24-2.01)	1.22 (0.80-1.86)
ATM	294	196	22	7	150	2.10 (1.71-2.57)	2.33 (1.87-2.91)	1.01 (0.64-1.59)	0.91 (0.42-1.95)	253	151	19	5	134	1.82 (1.46-2.27)	1.96 (1.52-2.53)	1.04 (0.59-1.72)	0.5 (0.12-1.36)	547	347	41	12	284	1.97 (1.69-2.28)	2.16 (1.83-2.56)	1.02 (0.72-1.45)	0.77 (0.40-1.47)
BARD1	62	24	27	12	32	2.09 (1.35-3.23)	1.40 (0.81-2.42)	5.99 (3.51-10.21)	9.29 (4.58-18.85)	49	20	11	6	35	1.37 (0.87-2.16)	0.91 (0.49-1.64)	2.52 (1.18-5.00)	3.18 (1.16-7.42)	111	44	38	18	67	1.71 (1.25-2.34)	1.15 (0.77-1.73)	4.41 (2.87-6.78)	6.26 (3.57-10.99)
RAD51C	54	24	20	10	26	1.93 (1.2-3.11)	1.31 (0.74-2.30)	3.99 (2.20-7.26)	5.71 (2.69-12.13)	41	16	9	4	35	1.20 (0.75-1.93)	0.83 (0.44-1.54)	2.19 (0.97-4.49)	**2.55 (0.90-7.17)**	95	40	29	14	61	1.52 (1.09-2.12)	1.07 (0.70-1.62)	3.18 (1.99-5.09)	**4.32 (2.35-7.94)**
RAD51D	51	26	13	9	25	1.80 (1.11-2.93)	1.52 (0.87-2.65)	2.92 (1.47-5.78)	6.01 (2.73-13.24)	26	13	7	1	14	1.72 (0.88-3.51)	1.61 (0.71-3.70)	3.93 (1.40-10.29)	**1.59 (0.21-12.09)**	77	39	20	10	39	1.77 (1.19-2.64)	1.55 (0.98-2.46)	3.21 (1.83-5.65)	**5.05 (2.42-10.53)**
TP53	7	3	2	0	2	3.06 (0.63-14.91)	1.95 (0.32-11.82)	5.42 (0.75-39.24)	0 (0-Inf)	19	9	2	2	2	9.59 (2.23-41.19)	**7.95 (1.72-36.80)**	**8.56 (1.20-60.77)**	**22.27 (3.14-158.24)**	26	12	4	2	4	5.67 (1.25-25.84)	4.41 (1.37-14.19)	**6.82 (1.70-27.47)**	**22.27 (3.14-158.24)**
PTEN	14	9	0	0	6	2.25 (0.85-6.00)	2.42 (0.84-6.97)	0 (0-Inf)	0 (0-Inf)	8	3	0	0	3	2.69 (0.71-10.15)	**1.77 (0.36-8.75)**	0 (0-Inf)	0 (0-Inf)	22	12	0	0	9	2.40 (1.09-5.26)	2.20 (0.91-5.32)	NA	NA
NF1	31	15	7	2	17	1.76 (0.96-3.21)	1.25 (0.61-2.55)	2.46 (1.01-6.02)	2.02 (0.46-8.82)	19	10	2	1	11	1.93 (0.91-4.31)	1.63 (0.65-4.03)	**1.56 (0.34-7.02)**	**2.02 (0.26-15.68)**	50	25	9	3	28	1.82 (1.13-2.94)	1.38 (0.79-2.43)	**2.18 (1.01-4.71)**	**2.02 (0.61-6.70)**
STK11	6	3	0	0	5	1.60 (0.48-5.28)	1.56 (0.35-7.03)	0 (0-Inf)	0 (0-Inf)	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA
CDH1	11	8	2	1	12	0.86 (0.37-1.98)	1.05 (0.42-2.63)	1.11 (0.24-5.10)	1.44 (0.18-11.28)	17	13	3	1	6	2.50 (1.01-7.07)	3.37 (1.24-10.72)	**4.28 (1.07-17.12)**	**3.71 (0.45-30.83)**	28	21	5	2	18	1.36 (0.72-2.56)	1.71 (0.85-3.45)	**2.33 (0.83-6.50)**	**2.29 (0.52-10.04)**

Supplementary Table 6: *CDH1* variant counts and odds ratios stratified by tumour histology. BC = breast cancer, UKB = UK Biobank, PTV = protein-truncating variant. A: Odds ratios (ORs) and 95% confidence intervals (CIs) for breast cancer for *CDH1* pathogenic variant carriers in UK Biobank study data, stratified by cancer histology of cases. B: Carrier counts and ORs towards lobular breast cancer for *CDH1* pathogenic variant carriers in UK Biobank, CARRIERS and a combined meta-analysis of both datasets. ORs in bold are deemed significant with lower 95% CI >1.

A

	Case cohort (all compared to UKB participants without BC history)		
Variant type	**All BC**	**Lobular BC**	**Non-lobular/ unknown BC**
PTVs	**3.83** (1.63-9.00)	**19.10** (7.15-51.05)	1.32 (0.31-5.56)
Non-PTVs	**5.57** (1.00-31.06)	**36.16** (6.44-202.91)	0.00094 ($2.06x10^{-34}$-$3.97x10^{27}$)
All	**4.11** (1.92-8.81)	**22.01** (9.45-51.31)	1.09 (0.26-4.59)

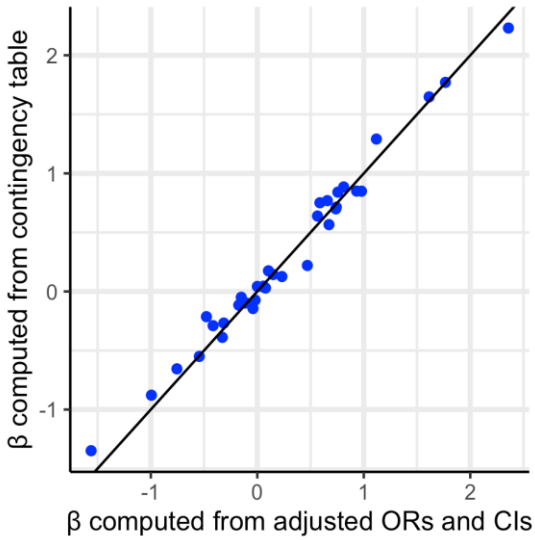
B

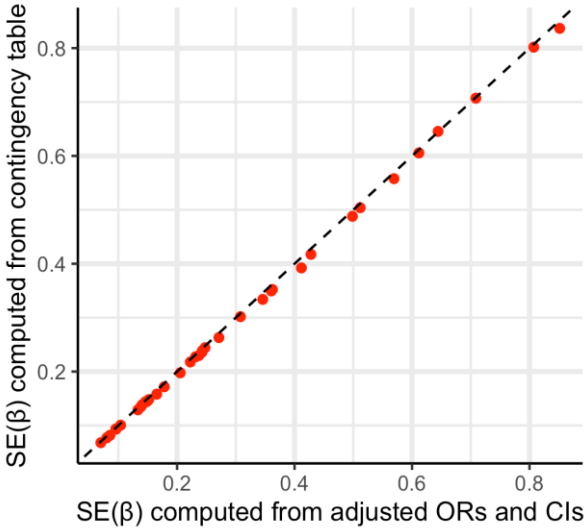
	Cases with variant	Controls with variant	Cases without variant	Controls without variant	OR (PTVs + missense)
CARRIERS	7	6	2992	32538	**15.74** (5.08-50.22)
UK Biobank	7	27	3101	229670	**20.61** (8.37-50.71)
Combined	14	33	6093	262208	**19.56** (9.90-38.62)

Supplementary Figures

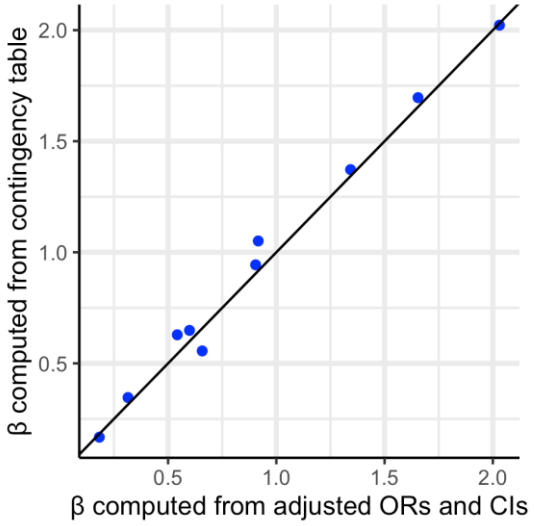
Supplementary Figure 1: Sensitivity analysis for logistic regression adjustment in BRIDGES and CARRIERS.
Comparison of unadjusted (computed directly from case and control variant counts) and logistic regression-adjusted (as published) betas (β) and standard errors (SE) for each gene surveyed in (A) BRIDGES and (B) CARRIERS. The difference between adjusted and unadjusted betas for a given gene in BRIDGES was added to the unadjusted betas calculated for BRIDGES pathogenic missense variants to simulate the predicted effect of logistic regression adjustment, as logistic regression could not be conducted for these variants independently (in the absence of participant-level data).

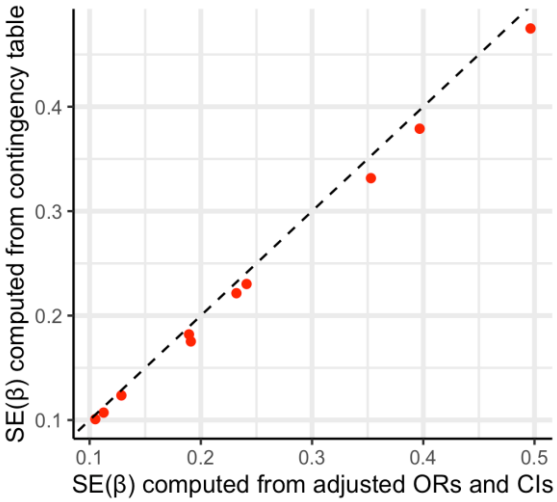
A





B





References

1. Yadav S, Hu C, Nathanson KL, et al. Germline Pathogenic Variants in Cancer Predisposition Genes Among Women With Invasive Lobular Carcinoma of the Breast. *Journal of Clinical Oncology* 2021; **39**(35): 3918-3926.
2. Easton DF, Pharoah PD, Antoniou AC, et al. Gene-panel sequencing and the prediction of breast-cancer risk. *The New England Journal of Medicine* 2015; **372**(23): 2243-57.