Supplementary Information

Article title: Physical activity, sedentary time and breast cancer risk: A Mendelian randomization study

Authors:

Suzanne C. Dixon-Suen, Sarah J. Lewis, Richard M. Martin, Dallas R. English, Terry Boyle, Graham G. Giles, Kyriaki Michailidou, Manjeet K. Bolla, Qin Wang, Joe Dennis, Michael Lush, ABCTB Investigators, Thomas U. Ahearn, Christine B. Ambrosone, Irene L. Andrulis, Hoda Anton-Culver, Volker Arndt, Kristan J. Aronson, Annelie Augustinsson, Päivi Auvinen, Laura E. Beane Freeman, Heiko Becher, Matthias W. Beckmann, Sabine Behrens, Marina Bermisheva, Carl Blomqvist, Natalia V. Bogdanova, Stig E. Bojesen, Bernardo Bonanni, Hermann Brenner, Thomas Brüning, Saundra S. Buys, Nicola J. Camp, Daniele Campa, Federico Canzian, Jose E. Castelao, Melissa H. Cessna, Jenny Chang-Claude, Stephen J. Chanock, Christine L. Clarke, Don M. Conroy, Fergus J. Couch, Angela Cox, Simon S. Cross, Kamila Czene, Mary B. Daly, Peter Devilee, Thilo Dörk, Miriam Dwek, Diana M. Eccles, A. Heather Eliassen, Christoph Engel, Mikael Eriksson, D. Gareth Evans, Peter A. Fasching, Olivia Fletcher, Henrik Flyger, Lin Fritschi, Marike Gabrielson, Manuela Gago-Dominguez, Montserrat García-Closas, José A. García-Sáenz, Mark S. Goldberg, Pascal Guénel, Melanie Gündert, Eric Hahnen, Christopher A. Haiman, Lothar Häberle, Niclas Håkansson, Per Hall, Ute Hamann, Steven N. Hart, Michelle Harvie, Peter Hillemanns, Antoinette Hollestelle, Maartje J. Hooning, Reiner Hoppe, John L. Hopper, Anthony Howell, David J. Hunter, Anna Jakubowska, Wolfgang Janni, Esther M. John, Audrey Jung, Rudolf Kaaks, Renske Keeman, Cari M. Kitahara, Stella Koutros, Peter Kraft, Vessela N. Kristensen, Katerina Kubelka-Sabit, Allison W. Kurian, James V. Lacey, Diether Lambrechts, Loic Le Marchand, Annika Lindblom, Sibylle Loibl, Jan Lubiński, Arto Mannermaa, Mehdi Manoochehri, Sara Margolin, Maria Elena Martinez, Dimitrios Mavroudis, Usha Menon, Anna Marie Mulligan, Rachel A. Murphy, NBCS Collaborators, Heli Nevanlinna, Ines Nevelsteen, William G. Newman, Kenneth Offit, Andrew F. Olshan, Håkan Olsson§, Nick Orr, Alpa V. Patel, Julian Peto, Dijana Plaseska-Karanfilska, Nadege Presneau, Brigitte Rack, Paolo Radice, Erika Rees-Punia, Gad Rennert, Hedy S. Rennert, Atocha Romero, Emmanouil Saloustros, Dale P. Sandler, Marjanka K. Schmidt, Rita K. Schmutzler, Lukas Schwentner, Christopher Scott, Mitul Shah,

Xiao-Ou Shu, Jacques Simard, Melissa C. Southey, Jennifer Stone, Harald Surowy, Anthony J. Swerdlow, Rulla M. Tamimi, William J. Tapper, Jack A. Taylor, Mary Beth Terry, Rob A.E.M. Tollenaar, Melissa A. Troester, Thérèse Truong, Michael Untch, Celine M. Vachon, Vijai Joseph, Barbara Wappenschmidt, Clarice R. Weinberg, Alicja Wolk, Drakoulis Yannoukakos, Wei Zheng, Argyrios Ziogas, Alison M. Dunning, Paul D.P. Pharoah, Douglas F. Easton, Roger L. Milne †, Brigid M. Lynch † *, on behalf of the Breast Cancer Association Consortium

§ deceased

* **Corresponding author**: Assoc. Prof. Brigid M. Lynch, Cancer Epidemiology Division, Cancer Council Victoria, Melbourne, Australia. Email: brigid.lynch@cancervic.org.au.

† Joint senior authors

Contents

Supplementary Methods

Mendelian randomization overview

Mendelian randomization (MR) is a form of instrumental variable analysis where exposures are measured indirectly using genotype. By measuring exposures using genetic proxies, which are randomised at meiosis (before conception) and are therefore less prone to bias such as selection bias, reverse causality, and confounding, MR may be able to provide estimates which more closely reflect underlying causal relationships. Since genotype is randomly allocated, MR studies simulate the design of a randomized controlled trial with groups defined by genotype analogous to trial arms, theoretically allowing MR to overcome key sources of bias in observational studies.

Defining genetic instruments: additional information

In defining instruments for our analysis, we selected genetic variants identified from GWAS models unadjusted for adiposity to avoid possible collider bias introduced by adjustment for a trait which is causally downstream.(1) We used the National Cancer Institute's LDpair or LDmatrix applications (2) to confirm that SNPs in each instrument were in linkage equilibrium (independent) (highest r^2=0.004). All SNPs were well-imputed (information score ≥0.90 for all but two SNPs; lowest score 0.84) (Table S2).

Overall physical activity: additional information

Our primary physical activity instrument was derived from a recent GWAS which used UK Biobank data on movement measured by wrist-worn triaxial accelerometers.(3) Working with the UK Biobank Accelerometer Working Group, Doherty and colleagues derived a measure for overall activity, assessed as average vector magnitude (in milligravities) per 30-second period, recorded across an accelerometer wear period of three to seven days.(3, 4) They identified five SNPs associated with this phenotype at conventional genome-wide significance ($p<5x10^{-8}$).(3) A separate research group performed a GWAS of multiple physical activity measures using UK Biobank data,(5) including the overall activity (average accelerations) phenotype derived by the UK Biobank Accelerometer Working Group.(4). Of the ten SNPs Klimentidis and colleagues identified as associated with overall activity at relaxed significance ($p<5x10^{-7}$), five signals overlapped with the Doherty-identified variants.(5, 6)

Vigorous physical activity: additional information

Klimentidis *et al* examined UK Biobank physical activity data from wrist-worn accelerometers (n~91,000) and self-report (n~377,000) to identify SNPs associated at stringent significance ($p<5x10^{-9}$) with vigorous activity.(5)

Outcomes: additional information

Ki-67 data to determine luminal A/B subtype was unavailable.

Statistical analysis: additional information

For the multi-SNP instruments, we used SNP-exposure and SNP-outcome beta coefficients and standard errors to estimate odds ratios and 95% confidence intervals of the effect of each trait on each outcome from inverse-variance weighted (IVW) MR, using a multiplicative random-effects model with simple weights (first-order term from delta expansion).(7) IVW-MR averages estimates of the causal effect across multiple SNPs, weighted by SNP-exposure beta coefficients, to derive a summary estimate.(7, 8) For the single-SNP instrument (accelerations >425 milligravities) we used the Wald (ratio) MR technique, dividing the SNP-outcome association (ZY) by the SNP-exposure association (ZX) to estimate the causal OR. The ratio estimate of the causal effect using a SNP 'k' is $\beta ZY_k/\beta ZX_k$. IVW-MR averages these Wald ratios across SNPs.

In sensitivity analyses, we applied weighted median MR(9) and MR-Egger(10), complementary methods which relax different MR assumptions. Weighted median MR allows up to half of the genetic instruments to be invalid; MR-Egger allows horizontal pleiotropy (although it has lower statistical power than IVW MR). We inspected causal estimates considering each SNP individually (inspecting scatter plots of SNP-exposure and SNP-outcome associations, and forest plots of SNP-specific causal effects). We also performed leave-one-out analyses (omitting one SNP each time) to further explore the robustness of our results to instrument composition.

Causal effects were estimated using the 'MendelianRandomization'(11) package and outlier detection was performed using the 'MR-PRESSO' package.(12) Analyses were conducted and reported with reference to MR guidelines.(1, 13)

Supplementary Tables

Table S1. Acronyms and study names of Breast Cancer Association Consortium studies in the analysis

Study acronym	Study name	Reference(s)
2SISTER *	The Two Sister Study	(14)
ABCFS	Australian Breast Cancer Family Study	(15)
ABCS	Amsterdam Breast Cancer Study	(16)
ABCTB	Australian Breast Cancer Tissue Bank	(17)
AHS	Agricultural Health Study	(18, 19)
BBCC	Bavarian Breast Cancer Cases and Controls	(20, 21)
BBCS	British Breast Cancer Study	(22, 23)
BCEES	Breast Cancer Employment and Environment Study	(24)
BCFR-NY *	New York Breast Cancer Family Registry	(25-27)
BCFR-PA *	Philadelphia Breast Cancer Family Registry	(25, 28)
BCFR-UTAH *	Utah Breast Cancer Family Registry	(25, 28)
BCINIS	Breast Cancer In Northern Israel Study	(29, 30)
BREOGAN	Breast Oncology Galicia Network	(31-35)
BSUCH	Breast Cancer Study of the University Clinic Heidelberg	(36)
CBCS	Canadian Breast Cancer Study	(37-40)
CCGP	Crete Cancer Genetics Program	--
CECILE	CECILE Breast Cancer Study	(41)
CGPS	Copenhagen General Population Study	(42)
CPSII	Cancer Prevention Study-II Nutrition Cohort	(43)
CTS	California Teachers Study	(44)
DIETCOMPLYF	DietCompLyf Breast Cancer Survival Study	(45)
EPIC	European Prospective Investigation into Cancer and Nutrition	(46)
ESTHER	ESTHER Breast Cancer Study	(47)
FHRISK *	Family History Risk Study	(48, 49)
GC-HBOC *	German Consortium for Hereditary Breast and Ovarian Cancer	(50-53)
GENICA	Gene Environment Interaction & Breast Cancer in Germany	(54, 55)
GEPARSIXTO	A randomized phase II trial investigating the addition of carboplatin to neoadjuvant therapy for triple-negative and HER2-positive early breast cancer	(56-59)
GESBC	Genetic Epidemiologic Study of Breast Cancer by Age 50	(60)
HABCS	Hannover Breast Cancer Study	(61)
HCSC	Hospital Clinico San Carlos	(62, 63)
HEBCS *	Helsinki Breast Cancer Study	(64-66)
HMBCS	Hannover-Minsk Breast Cancer Study	(67)
HUBCS	Hannover-Ufa Breast Cancer Study	(67)
KARBAC *	Karolinska Breast Cancer Study	(68, 69)
KARMA	Karolinska Mammography Project for Risk Prediction of Breast Cancer – Cohort Study	(70)
KBCP	Kuopio Breast Cancer Project	(71, 72)
LMBC	Leuven Multidisciplinary Breast Centre	(73, 74)
MABCS	Macedonian Breast Cancer Study	--
MARIE	Mammary Carcinoma Risk Factor Investigation	(75)
MBCSG *	Milan Breast Cancer Study Group	(76, 77)
MCBCS	Mayo Clinic Breast Cancer Study	(78)
MCCS	Melbourne Collaborative Cohort Study	(79)
MEC	Multiethnic Cohort	(80)
MISS	Melanoma Inquiry of Southern Sweden	(81, 82)
MMHS	Mayo Mammography Health Study	(83)
MSKCC *	Memorial SloanKettering Cancer Center Study	(84)
MTLGEBCS	Montreal Gene-Environment Breast Cancer Study	--
NBCS	Norwegian Breast Cancer Study	(85-88)

Study acronym	Study name	Reference(s)
NBHS	Nashville Breast Health Study	(89)
NC-BCFR *	Northern California Breast Cancer Family Registry	(25, 28)
NCBCS	North Carolina Breast Cancer study	(90, 91)
NHS	Nurses' Health Study	(92, 93)
NHS2	Nurses' Health Study 2	(94)
OFBCR *	Ontario Familial Breast Cancer Registry	(25)
ORIGO	Leiden University Medical Centre Breast Cancer Study	(95, 96)
PBCS	NCI Polish Breast Cancer Study	(97)
pKARMA	Karolinska Mammography Project for Risk Prediction of Breast Cancer – Case-Control Study	--
PLCO	The Prostate, Lung, Colorectal and Ovarian (PLCO) Cancer Screening Trial	(98)
POSH	Prospective Study of Outcomes in Sporadic Versus Hereditary Breast Cancer	(99-104)
PREFACE	Evaluation of Predictive Factors regarding the Effectivity of Aromatase Inhibitor Therapy	(105)
PROCAS	Predicting the Risk Of Cancer At Screening Study	(48)
RBCS *	Rotterdam Breast Cancer Study	(106)
SBCS	Sheffield Breast Cancer Study	(107, 108)
SEARCH	Study of Epidemiology and Risk Factors in Cancer Heredity	(109)
SISTER *	The Sister Study	(110-112)
SKKDKFZS	Städtisches Klinikum Karlsruhe Deutsches Krebsforschungszentrum Study	(113)
SMC	Swedish Mammography Cohort	(114)
SUCCESSB	Simultaneous Study of Gemcitabine-Docetaxel Combination adjuvant treatment	(115-117)
SUCCESSC	Simultaneous Study of Docetaxel Based Anthracycline Free Adjuvant Treatment Evaluation	(118-121)
SZBCS	IHCC-Szczecin Breast Cancer Study	(122-125)
TNBCC	Triple Negative Breast Cancer Consortium Study	--
UBCS *	Utah Breast Cancer Study	(126, 127)
UCIBCS	UCI Breast Cancer Study	(128, 129)
UKBGS	UK Breakthrough Generations Study	(130)
UKOPS	UK Ovarian Cancer Population Study †	(131)
USRT	US Radiologic Technologists Study	(132-135)

-- No citation

* Included familial cases or controls (recruited on the basis of being at high risk for breast cancer).

† This study contributed controls only.

Table S2. Single nucleotide polymorphisms used as instruments for physical activity or sedentary time

SNP*	Chr	Position†	Function	Nearest gene	Effect‡ / Other allele	Beta for association with trait§	Standard error for association with trait§	UK Biobank effect allele‡ frequency	OncoArray effect allele‡ frequency	OncoArray imputation information score
Overall (average) activity: 5 SNPs associated at $p < 5 \times 10^{-8}$ (identified by Doherty *et al* (3))										
rs6775319	3	18,758,501	Intron	*SATB1-AS1*	A/T	0.027	0.005	0.27	0.30	0.99
rs6895232	5	152,039,421	--	--	T/A	0.027	0.005	0.66	0.69	0.99
rs564819152	10	21,820,650	Intron	*MLLT10*	A/G	0.028	0.005	0.68	0.65	0.99
rs2696625	17	44,326,864	Downstream	*MAPK8IP1P1*	G/A	0.037	0.005	0.23	0.21	0.97
rs59499656	18	40,768,309	Intergenic	*RIT2*	T/A	0.028	0.005	0.35	0.34	0.99
Overall (average) activity, secondary instrument: 10 SNPs associated at $p < 5 \times 10^{-7}$ (identified by Klimentidis *et al* (5, 6))**										
rs12045968	1	33,690,698	Intergenic	*ZNF362*	G/T	0.029	0.005	0.22	0.21	1.00
rs34517439	1	78,450,517	Intron	*DNAJB4*	C/A	0.038	0.007	0.91	0.92	1.00
rs6775319	3	18,758,501	Intron	*SATB1-AS1*	A/T	0.028	0.005	0.30	0.30	0.99
rs9293503	5	87,948,962	Intron	*LINC00461*	T/C	0.040	0.007	0.88	0.86	0.93
rs12522261	5	152,054,825	Intron	*LINC01470*	G/A	0.026	0.005	0.67	0.69	0.99
rs11012732	10	21,830,104	Intron	*MLLT10*	A/G	0.028	0.005	0.65	0.63	1.00
rs148193266	11	104,528,681	Intergenic	*RP11-681H10.1*	C/A	0.063	0.011	0.02	0.03	0.87
rs1550435	15	74,331,385	Intron	*PML*	T/C	0.025	0.005	0.53	0.54	0.99
rs55657917	17	43,844,560	--	--	G/T	0.036	0.005	0.22	0.20	1.00
rs59499656	18	40,768,309	Intergenic	*RIT2*	T/A	0.028	0.005	0.34	0.34	0.99
Fraction of time with accelerations >425 mg: 1 SNP associated at $p < 5 \times 10^{-9}$ (identified by Klimentidis *et al* (5))										
rs743580	15	74,328,116	Missense	*PML*	A/G ¶	0.025	0.00005	0.51	0.49	0.98
Self-reported vigorous physical activity: 5 SNPs associated at $p < 5 \times 10^{-9}$ (identified by Klimentidis *et al* (5))										
rs2764261	6	108,927,842	Intron	*FOXO3*	A/G	0.039	0.001	0.37	0.40	1.00
rs328902	7	35,020,843	Intron	*DPY19L1*	T/C	0.041	0.001	0.31	0.31	1.00
rs13243553	7	133,506,955	Intron	*EXOC4*	G/A	0.039	0.001	0.61	0.62	0.98
rs3781411	10	126,715,436	Missense	*CTBP2*	C/T	0.058	0.001	0.88	0.85	1.00
rs1248860	3	85,015,779	Intron	*CADM2*	A/G	0.041	0.001	0.52	0.51	0.99
Percent time spent sedentary: 6 SNPs associated at $p < 5 \times 10^{-8}$ (identified by Doherty *et al* (3))										
rs61776614	1	2,166,406	Intron	*SKI*	C/T	0.050	0.009	0.93	0.93	0.84
rs1858242	3	68,527,135	Intron	*FAM19A1*	A/G	0.031	0.005	0.26	0.25	0.99
rs26579	5	87,985,295	Intron	*LINC00461*	G/C	0.028	0.005	0.42	0.46	0.95
rs25981	5	106,822,908	Intron	*EFNA5*	G/C	0.028	0.005	0.53	0.53	0.99

SNP*	Chr	Position†	Function	Nearest gene	Effect‡ / Other allele	Beta for association with trait§	Standard error for association with trait§	UK Biobank effect allele‡ frequency	OncoArray effect allele‡ frequency	OncoArray imputation information score
rs6870096	5	151,945,811	Intergenic	*CTB-95D12.1*	G/C	0.028	0.005	0.68	0.69	0.98
rs34858520	7	71,723,883	Intron	*CALN1*	A/G	0.028	0.005	0.56	0.57	0.99

Abbreviations: Chr, chromosome; GWAS, genome-wide association study; SNP, single nucleotide polymorphism.

* the National Cancer Institute's LDpair or LDmatrix (2) applications were used to confirm that SNPs on the same chromosome within each instrument are independent (highest r^2=0.004).

† human genome assembly GRCh37 (hg19)

‡ Allele associated with an increase in the trait (i.e., with increased physical activity [physical activity instruments], or with increased time spent sedentary [sedentary time instrument])

§ Betas and standard errors for associations between SNPs and exposure (physical activity or sedentary time) are from, or derived from, the GWAS which identified the SNPs.

** Five signals overlap with the Doherty-identified variants for overall activity.

¶ This SNP (rs743580, A/G) has an effect allele frequency near 50% and the minor allele in UK Biobank (G) differs from that in OncoArray (A) , but it is not palindromic so the trait-increasing allele was easily identifiable in OncoArray data. Additionally, we confirmed that the trait-increasing allele, A, was positively associated with strenuous activity in BCAC.

Table S3. Comparison of results from different Mendelian randomization methods: Association between the primary instrumental genetic variables for overall physical activity (per standard deviation) and risk of breast cancer

Type of breast cancer	N cases (vs. 54,452 controls)	Odds ratios (95% CI) *†	P for heterogeneity‡ or pleiotropy§
Invasive cancers			
All invasive	69,838		
IVW		0.48 (0.30-0.78)	0.016
Weighted median		0.59 (0.39-0.90)	
MR-Egger		0.14 (0.00-11.0)	0.569
Pre/perimenopausal	¶ 23,999		
IVW		0.51 (0.31-0.83)	0.419
Weighted median		0.56 (0.30-1.07)	
MR-Egger		0.03 (0.00-1.50)	0.148
Postmenopausal	**45,839		
IVW		0.48 (0.28-0.80)	0.054
Weighted median		0.58 (0.36-0.94)	
MR-Egger		0.36 (0.00-50.7)	0.911
By receptor status			
ER+	46,528		
IVW		0.45 (0.25-0.83)	0.004
Weighted median		0.58 (0.37-0.91)	
MR-Egger		0.17 (0.00-48.1)	0.726
ER-	11,246		
IVW		0.79 (0.37-1.66)	0.069
Weighted median		0.66 (0.32-1.37)	
MR-Egger		0.45 (0.00-546)	0.875
PR+	34,891		
IVW		0.43 (0.22-0.85)	0.003
Weighted median		0.56 (0.33-0.94)	
MR-Egger		0.08 (0.00-40.5)	0.601
PR-	16,432		
IVW		0.65 (0.38-1.13)	0.186
Weighted median		0.63 (0.34-1.14)	
MR-Egger		0.76 (0.00-140)	0.953
HER2+	6,945		
IVW		0.48 (0.26-0.89)	0.479
Weighted median		0.47 (0.21-1.05)	
MR-Egger		0.01 (0.00-1.91)	0.149
HER2-	33,214		
IVW		0.58 (0.35-0.98)	0.060
Weighted median		0.64 (0.39-1.04)	
MR-Egger		0.17 (0.00-20.4)	0.613
Combined hormone receptor- and/or HER2-defined subtypes			
ER+ or PR+; HER2+	4,816		
IVW		0.42 (0.20-0.88)	0.478
Weighted median		0.57 (0.22-1.46)	
MR-Egger		0.00 (0.00-0.94)	0.087
ER+ or PR+; HER2-	27,874		
IVW		0.57 (0.28-1.18)	0.004
Weighted median		0.64 (0.37-1.09)	
MR-Egger		0.13 (0.00-106)	0.667

Type of breast cancer	N cases (vs. 54,452 controls)	Odds ratios (95% CI) *†	P for heterogeneity‡ or pleiotropy§
ER-; PR-; HER2+	1,974		
IVW		0.53 (0.18-1.57)	0.700
Weighted median		0.42 (0.11-1.68)	
MR-Egger		0.09 (0.00-801)	0.701
ER-; PR-; HER2-	4,964		
IVW		0.60 (0.17-2.12)	0.015
Weighted median		0.56 (0.20-1.59)	
MR-Egger		0.32 (0.00-51,961)	0.917
ER- and PR- (all)	9,215		
IVW		0.65 (0.27-1.56)	0.036
Weighted median		0.47 (0.21-1.07)	
MR-Egger		0.20 (0.00-841)	0.783
By morphology			
Ductal	42,223		
IVW		0.52 (0.32-0.84)	0.053
Weighted median		0.64 (0.41-1.02)	
MR-Egger		0.10 (0.00-7.61)	0.463
Lobular	8,795		
IVW		0.32 (0.18-0.58)	0.500
Weighted median		0.31 (0.14-0.68)	
MR-Egger		4.01 (0.03-533)	0.310
By stage at diagnosis			
Stage I	17,583		
IVW		0.51 (0.32-0.82)	0.333
Weighted median		0.47 (0.26-0.85)	
MR-Egger		0.11 (0.00-7.18)	0.471
Stage II	15,992		
IVW		0.36 (0.22-0.58)	0.576
Weighted median		0.35 (0.18-0.66)	
MR-Egger		0.11 (0.00-5.81)	0.553
Stage III/IV	4,553		
IVW		0.37 (0.17-0.81)	0.499
Weighted median		0.34 (0.13-0.94)	
MR-Egger		0.10 (0.00-70.0)	0.687
By tumor grade			
Grade 1/2	34,647		
IVW		0.43 (0.23-0.81)	0.011
Weighted median		0.54 (0.33-0.89)	
MR-Egger		0.18 (0.00-63.2)	0.768
Grade 3	16,432		
IVW		0.46 (0.30-0.72)	0.552
Weighted median		0.42 (0.23-0.75)	
MR-Egger		0.28 (0.01-10.7)	0.786
In situ cancers			
All in situ	6,667		
IVW		0.63 (0.34-1.18)	0.390
Weighted median		0.71 (0.32-1.59)	
MR-Egger		0.01 (0.00-1.24)	0.087
Ductal carcinoma in situ	3,510		
IVW		0.92 (0.25-3.43)	0.039
Weighted median		1.01 (0.31-3.25)	
MR-Egger		0.00 (0.00-0.12)	0.011

Abbreviations: CI, confidence interval; ER+/-, estrogen receptor positive/negative; GWAS, genome wide association study; HER2+/-, human epidermal growth factor receptor 2 positive/negative; IVW, inverse-variance weighted; MR, Mendelian Randomization; PR+/-, progesterone receptor positive/negative; SNP, single nucleotide polymorphism.

* Causal odds ratios were estimated by three different Mendelian randomization methods, using five SNPs identified in a GWAS of accelerometer-measured movement traits by Doherty et al (3)

† Confidence intervals from weighted median MR were generated from 10,000 bootstrap samples (seed 314159265)

‡ Relating to IVW MR estimates: p-value associated with the heterogeneity test statistic (Cochran's Q statistic) measuring heterogeneity of causal effects between SNPs

§ Relating to MR-Egger estimates: p-value for the intercept test (p-value for pleiotropy)

¶ vs pre/perimenopausal controls (n=17,686), assigned using age (<50 years) if menopause status was unknown

** vs postmenopausal controls (n=36,766), assigned using age (≥50 years) if menopause status was unknown

Table S4. Leave-one-out analyses: Association between the primary instrumental genetic variables for overall physical activity (per standard deviation) and risk of breast cancer, omitting one SNP at a time

	Full instrument (five SNPs)		Excluding rs6775319 *		Excluding rs6895232 **		Excluding rs564819152 †		Excluding rs2696625 ††		Excluding rs59499656 ‡	
Type of breast cancer	**Odds ratios (95% CI) §**	**P_{het}¶**	**Odds ratios (95% CI) §**	**P_{het}¶**	**Odds ratios (95% CI) §**	**P_{het}¶**	**Odds ratios (95% CI) §**	**P_{het}¶**	**Odds ratios (95% CI) §**	**P_{het}¶**	**Odds ratios (95% CI) §**	**P_{het}¶**
Invasive cancers												
All invasive	0.48 (0.30-0.78)	0.016	0.46 (0.25-0.82)	0.010	0.44 (0.25-0.79)	0.014	0.59 (0.42-0.83)	0.312	0.52 (0.28-0.97)	0.010	0.43 (0.25-0.72)	0.031
Pre/perimenopausal	0.51 (0.31-0.83)	0.419	0.46 (0.27-0.78)	0.402	0.48 (0.27-0.87)	0.306	0.57 (0.33-0.99)	0.383	0.63 (0.36-1.10)	0.565	0.45 (0.26-0.78)	0.404
Postmenopausal	0.48 (0.28-0.80)	0.054	0.45 (0.24-0.86)	0.030	0.44 (0.24-0.82)	0.040	0.61 (0.42-0.89)	0.816	0.48 (0.24-0.95)	0.025	0.42 (0.23-0.77)	0.058
By receptor status												
ER+	0.45 (0.25-0.83)	0.004	0.41 (0.20-0.84)	0.004	0.40 (0.20-0.80)	0.006	0.60 (0.43-0.85)	0.459	0.47 (0.21-1.05)	0.002	0.42 (0.20-0.88)	0.003
ER-	0.79 (0.37-1.66)	0.069	0.96 (0.44-2.06)	0.122	0.90 (0.38-2.18)	0.059	0.60 (0.31-1.17)	0.247	0.87 (0.33-2.31)	0.041	0.67 (0.28-1.60)	0.068
PR+	0.43 (0.22-0.85)	0.003	0.40 (0.18-0.92)	0.002	0.38 (0.17-0.85)	0.003	0.58 (0.37-0.91)	0.223	0.48 (0.20-1.14)	0.002	0.36 (0.17-0.76)	0.010
PR-	0.65 (0.38-1.13)	0.186	0.68 (0.35-1.34)	0.111	0.80 (0.49-1.30)	0.472	0.53 (0.33-0.87)	0.412	0.67 (0.33-1.39)	0.105	0.61 (0.31-1.19)	0.125
HER2+	0.48 (0.26-0.89)	0.479	0.41 (0.21-0.82)	0.475	0.47 (0.22-0.98)	0.323	0.52 (0.26-1.06)	0.364	0.62 (0.30-1.27)	0.685	0.40 (0.20-0.80)	0.502
HER2-	0.58 (0.35-0.98)	0.060	0.57 (0.30-1.09)	0.030	0.49 (0.30-0.81)	0.171	0.72 (0.49-1.07)	0.385	0.62 (0.31-1.23)	0.033	0.54 (0.29-1.03)	0.039
Combined hormone receptor- and/or HER2-defined subtypes												
ER+/PR+; HER2+	0.42 (0.20-0.88)	0.478	0.36 (0.16-0.81)	0.448	0.38 (0.17-0.85)	0.380	0.45 (0.19-1.07)	0.336	0.61 (0.26-1.41)	0.852	0.38 (0.17-0.86)	0.380
ER+/PR+; HER2-	0.57 (0.28-1.18)	0.004	0.52 (0.22-1.22)	0.003	0.47 (0.23-0.97)	0.020	0.79 (0.49-1.26)	0.254	0.61 (0.24-1.58)	0.002	0.55 (0.22-1.37)	0.002
ER-; PR-; HER2+	0.53 (0.18-1.57)	0.700	0.41 (0.13-1.36)	0.756	0.69 (0.21-2.30)	0.758	0.56 (0.16-1.87)	0.539	0.61 (0.17-2.14)	0.569	0.44 (0.13-1.48)	0.628
ER-; PR-; HER2-	0.60 (0.17-2.12)	0.015	0.95 (0.37-2.44)	0.224	0.61 (0.12-3.04)	0.007	0.39 (0.12-1.25)	0.094	0.69 (0.13-3.63)	0.008	0.51 (0.11-2.46)	0.009
ER- and PR- (all)	0.65 (0.27-1.56)	0.036	0.81 (0.32-2.03)	0.070	0.74 (0.26-2.15)	0.027	0.46 (0.22-0.96)	0.226	0.75 (0.24-2.33)	0.024	0.54 (0.19-1.54)	0.034
By morphology												
Ductal	0.52 (0.32-0.84)	0.053	0.47 (0.27-0.84)	0.045	0.48 (0.27-0.85)	0.041	0.63 (0.43-0.91)	0.346	0.57 (0.31-1.05)	0.039	0.46 (0.27-0.79)	0.067
Lobular	0.32 (0.18-0.58)	0.500	0.33 (0.17-0.65)	0.348	0.36 (0.19-0.69)	0.435	0.39 (0.20-0.74)	0.581	0.27 (0.14-0.54)	0.488	0.27 (0.14-0.53)	0.550
By stage at diagnosis												
Stage I	0.51 (0.32-0.82)	0.333	0.46 (0.28-0.75)	0.357	0.51 (0.28-0.94)	0.205	0.60 (0.37-0.98)	0.488	0.56 (0.31-1.01)	0.259	0.44 (0.27-0.72)	0.400
Stage II	0.36 (0.22-0.58)	0.576	0.36 (0.21-0.61)	0.408	0.31 (0.19-0.53)	0.693	0.42 (0.25-0.71)	0.718	0.39 (0.22-0.67)	0.446	0.33 (0.20-0.57)	0.473
Stage III/IV	0.37 (0.17-0.81)	0.499	0.46 (0.20-1.06)	0.570	0.28 (0.12-0.65)	0.854	0.38 (0.15-0.94)	0.340	0.41 (0.16-1.03)	0.360	0.37 (0.15-0.92)	0.339
By tumor grade												
Grade 1/2	0.43 (0.23-0.81)	0.011	0.38 (0.19-0.73)	0.023	0.40 (0.19-0.85)	0.007	0.58 (0.39-0.85)	0.514	0.45 (0.20-1.02)	0.005	0.40 (0.19-0.87)	0.007
Grade 3	0.46 (0.30-0.72)	0.552	0.51 (0.32-0.82)	0.546	0.43 (0.27-0.70)	0.466	0.50 (0.31-0.82)	0.477	0.48 (0.29-0.81)	0.402	0.40 (0.25-0.65)	0.722
In situ cancers												
All in situ	0.63 (0.34-1.18)	0.390	0.53 (0.27-1.04)	0.485	0.57 (0.27-1.22)	0.299	0.59 (0.27-1.28)	0.274	0.85 (0.42-1.74)	0.666	0.69 (0.32-1.50)	0.286
DCIS	0.92 (0.25-3.43)	0.039	0.69 (0.16-2.90)	0.055	0.65 (0.16-2.64)	0.071	0.82 (0.15-4.32)	0.021	1.69 (0.53-5.36)	0.228	1.16 (0.24-5.68)	0.031

Abbreviations: CI, confidence interval; DCIS, ductal carcinoma in situ; ER+/-, estrogen receptor positive/negative; GWAS, genome wide association study; HER2+/-, human epidermal growth factor receptor 2 positive/negative; PR+/-, progesterone receptor positive/negative; SNP, single nucleotide polymorphism.

* This SNP was identified by MR-PRESSO as an outlier for analyses of triple negative cancers (ER-/PR-/HER2-). It was also associated with adiposity in a prior GWAS.

** This SNP is correlated with a SNP predicting sedentary behaviour, rs6870096 (r^2=0.25 using the National Cancer Institute's LDpair application(2)).

† This SNP was identified by pleiotropy investigations as an outlier for analyses of all invasive, ER+, PR+, HR+/HER2-, HR-, and well/moderately differentiated cancers. This SNP was associated with ovarian cancer risk in a prior GWAS.

†† This SNP was associated with ovarian cancer risk in a prior GWAS.

‡ This SNP was associated with adiposity in a prior GWAS.

§ Causal odds ratios were estimated by inverse-variance weighted Mendelian randomization, using SNPs identified in a GWAS of accelerometer-measured movement traits by Doherty et al (3)

¶ p-value associated with the heterogeneity test statistic (Cochran's Q statistic) measuring heterogeneity of causal effects between SNPs

Table S5. Other phenotypes or gene expression differences associated with single nucleotide polymorphisms used in analysis as instruments for physical activity or sedentary time

			Characteristics associated with analysed (or correlated, $r^2 \geq 0.8$) SNPs in prior genome-wide association or gene expression studies (at $p < 5 \times 10^{-8}$), by direction of effect for allele predicting greater activity (PA instruments) or more sedentary time (sedentary time instrument) †				Relevance for our study
SNP	**Chr**	**Position***	**Traits: Conceivable confounders**	**Traits: Adiposity-related (conceivable mediators)**	**Traits: Cancer risk**	**Gene expression changes in breast tissue**	
Overall (average) activity: 5 SNPs associated at $p < 5 \times 10^{-8}$ (identified by Doherty *et al* (3))							
rs6775319	3	18,758,501	--	↓ Fat percentage (136)	--	--	Effect of PA on BC risk may be partially mediated through reduced adiposity
rs6895232	5	152,039,421	--	--	--	--	--
rs564819152	10	21,820,650	--	--	↓ Ovarian cancer (137)	--	Possible reflection of a confounding effect (away from null) along tumorigenic pathways; Reported results omitting this SNP
rs2696625	17	44,326,864	--	--	↑ Ovarian cancer (137)	--	Conceivable reflection of a confounding effect, but the association would likely bias toward the null; Reported results omitting this SNP
rs59499656	18	40,768,309	--	↓: Fat mass, Fat percentage, Waist circumference, Weight (136)	--	--	Effect of PA on BC risk may be partially mediated through reduced adiposity
Overall (average) activity, secondary instrument: 10 SNPs associated at $p < 5 \times 10^{-7}$ (identified by Klimentidis *et al* (5, 6))							
rs12045968	1	33,690,698	--	--	--	--	--
rs34517439	1	78,450,517	↓: Height (136); Psoriasis (138)	↓: Weight, Hip circumference, Fat mass, Basal metabolic rate, BMI, Waist circumference, Fat percentage (136)	↓ Lung cancer (139)	--	Any effect of PA on BC risk may be partially mediated through reduced adiposity; Possible confounding (height, psoriasis, unmeasured

SNP	Chr	Position*	Characteristics associated with analysed (or correlated, $r^2 \geq 0.8$) SNPs in prior genome-wide association or gene expression studies (at $p < 5x10^{-8}$), by direction of effect for allele predicting greater activity (PA instruments) or more sedentary time (sedentary time instrument) †				Relevance for our study
			Traits: Conceivable confounders	**Traits: Adiposity-related (conceivable mediators)**	**Traits: Cancer risk**	**Gene expression changes in breast tissue**	
							tumorigenic processes); Reported results omitting this SNP
rs6775319	3	18,758,501	--	↓ Fat percentage (136)	--	--	Effect of PA on BC risk may be partially mediated through reduced adiposity
rs9293503	5	87,948,962	--	--	--	--	--
rs12522261	5	152,054,825	--	--	--	--	--
rs11012732	10	21,830,104	--	↓: Fat percentage, Weight, Fat mass, Waist circumference, Hip circumference, BMI (136)	↑ Meningioma (140) ↓ Ovarian cancer (137)	--	Any effect of PA on BC risk may be partially mediated through reduced adiposity; Possible reflection of confounding effects along tumorigenic pathways; Reported results omitting this SNP
rs148193266	11	104,528,681	--	--	--	--	--
rs1550435	15	74,331,385	↑ Height (136, 141)	--	--	--	Possible confounding, but the association would likely bias toward the null; Reported results omitting this SNP
rs55657917	17	43,844,560	↑: Alcohol intake frequency, Medication for pain relief/ constipation/heartburn (136) ↓: Height, College qualifications, Daytime dozing or sleeping/ napping (136)	--	↑ Ovarian cancer (137)	↑ Expression of nearby genes *ARL17A*, *CRHR1*, *CRHR1-IT1*, *DND1P1*, *KANSL1-AS1*, *LRRC37A*, *LRRC37A2*, *RPS26P8*, collectively associated at $p < 5x10^{-8}$ with >150 traits including alcohol intake frequency, bone mineral density, **breast cancer in BRCA1 and BRCA2**	Possible confounding (via influencing alcohol intake, medication use, height, education, unmeasured tumorigenic processes, or via altering gene expression levels of genes associated with confounders, or **directly with breast cancer**); Reported results omitting this SNP

SNP	Chr	Position*	Characteristics associated with analysed (or correlated, $r^2 \geq 0.8$) SNPs in prior genome-wide association or gene expression studies (at $p<5x10^{-8}$), by direction of effect for allele predicting greater activity (PA instruments) or more sedentary time (sedentary time instrument) †				Relevance for our study
			Traits: Conceivable confounders	**Traits: Adiposity-related (conceivable mediators)**	**Traits: Cancer risk**	**Gene expression changes in breast tissue**	
						carriers, education, fat mass/%, forced expiratory volume, height, and ovarian cancer ↓ Expression of nearby gene *LRRC37A4P* (associated at $p<5x10^{-8}$ with 18 traits, primarily red and white blood cell characteristics)	
rs59499656	18	40,768,309	--	↓: Fat mass, Fat percentage, Waist circumference, Weight (136)	--	--	Effect of PA on BC risk may be partially mediated through reduced adiposity
Fraction of time with accelerations >425 mg: 1 SNP associated at $p < 5 \times 10^{-9}$ (identified by Klimentidis *et al* (5))							
rs743580	15	74,328,116	↑ Height (136, 142)	↓: Fat percentage, BMI (136)	--	--	Any effect of PA on BC risk may be partially mediated through reduced adiposity; Conceivable confounding, but the association would likely bias toward the null; Reported results omitting this SNP
Self-reported vigorous physical activity: 5 SNPs associated at $p < 5 \times 10^{-9}$ (identified by Klimentidis *et al* (5))							
rs2764261	6	108,927,842	↑ Age at menarche (136) ↓ Height (136)	↓: Weight, Fat mass, Hip circumference, Basal metabolic rate, Waist circumference, BMI, Fat percentage (136)	--	--	Any effect of PA on BC risk may be partially mediated through reduced adiposity; Possible confounding (age at menarche, height); Reported results omitting this SNP
rs328902	7	35,020,843	--	--	--	--	--
rs13243553	7	133,506,955	--	--	--	--	--

			Characteristics associated with analysed (or correlated, $r^2 \geq 0.8$) SNPs in prior genome-wide association or gene expression studies (at $p < 5x10^{-8}$), by direction of effect for allele predicting greater activity (PA instruments) or more sedentary time (sedentary time instrument) †				Relevance for our study
SNP	**Chr**	**Position***	**Traits: Conceivable confounders**	**Traits: Adiposity-related (conceivable mediators)**	**Traits: Cancer risk**	**Gene expression changes in breast tissue**	
rs3781411	10	126,715,436	--	--	--	--	--
rs1248860	3	85,015,779	↑ Comparative body size age 10 (136) ↓ Past tobacco smoking (136)	--	--	--	Possible confounding (early-life body size, smoking); Reported results omitting this SNP
Percent time spent sedentary: 6 SNPs associated at $p < 5 \times 10^{-8}$ (identified by Doherty *et al* (3))							
rs61776614	1	2,166,406	--	--	--	--	--
rs1858242	3	68,527,135	--	--	--	--	--
rs26579	5	87,985,295	↑ Years of education (143); College Qualifications, High-school completion qualifications, Other professional qualifications (136)	↑ Trunk fat percentage (136)	--	--	Effect of sedentary behaviour on BC risk may be partially mediated through increased adiposity; Possible confounding (education); Reported results omitting this SNP
rs25981	5	106,822,908	--	--	--	--	--
rs6870096	5	151,945,811	--	--	--	--	--
rs34858520	7	71,723,883	--	--	--	--	--

Abbreviations: BC, breast cancer; BMI, body mass index; Chr, chromosome; DIY, do-it-yourself; GWAS, genome-wide association study; PA, physical activity; SNP, single nucleotide polymorphism; WHR, waist-hip ratio.

* human genome assembly GRCh37 (hg19)

† Data from the University of Cambridge PhenoScanner V2(144, 145) or NHGRI-EBI GWAS Catalog (146) (as of October 2020); arrows denote direction of effect (risk association or gene expression change) relating to the allele which is also associated with increased physical activity (activity instruments) or increased sedentary time (sedentary behaviour instrument). Expression data was from Genotype-Tissue Expression (GTEx) project.(147)

Table S6. Association between the secondary instrumental genetic variables for overall physical activity (per standard deviation) and risk of breast cancer

		Full instrument (ten SNPs)		Excluding one pleiotropic SNP for outcomes with detected pleiotropy *	
Type of breast cancer	**N cases (vs. 54,452 controls)**	**Odds ratios (95% CI) †**	**P for heterogeneity‡**	**Odds ratios (95% CI) †**	**P for heterogeneity‡**
Invasive cancers					
All invasive	69,838	0.61 (0.44-0.85)	0.010	0.71 (0.57-0.88)	0.596
Pre/perimenopausal	§ 23,999	0.65 (0.45-0.93)	0.494	--	
Postmenopausal	¶ 45,839	0.60 (0.43-0.84)	0.068	--	
By receptor status					
ER+	46,528	0.57 (0.39-0.83)	0.004	0.69 (0.54-0.88)	0.931
ER-	11,246	0.72 (0.45-1.17)	0.109	--	
PR+	34,891	0.55 (0.36-0.84)	0.003	0.67 (0.51-0.88)	0.647
PR-	16,432	0.66 (0.48-0.92)	0.443	--	
HER2+	6,945	0.58 (0.34-0.97)	0.254	--	
HER2-	33,214	0.70 (0.50-0.99)	0.072	--	
Combined hormone receptor- and/or HER2-defined subtypes					
ER+ or PR+; HER2+	4,816	0.49 (0.29-0.85)	0.458	--	
ER+ or PR+; HER2-	27,874	0.68 (0.45-1.04)	0.010	0.83 (0.62-1.12)	0.729
ER-; PR-; HER2+	1,974	0.74 (0.33-1.68)	0.468	--	
ER-; PR-; HER2-	4,964	0.73 (0.36-1.47)	0.088	--	
ER- and PR- (all)	9,215	0.70 (0.42-1.17)	0.126	--	
By morphology					
Ductal	42,223	0.63 (0.46-0.86)	0.067	--	
Lobular	8,795	0.45 (0.28-0.71)	0.358	--	
By stage at diagnosis					
Stage I	17,583	0.59 (0.43-0.82)	0.558	--	
Stage II	15,992	0.52 (0.35-0.78)	0.231	--	
Stage III/IV	4,553	0.51 (0.28-0.93)	0.385	--	
By tumor grade					
Grade 1/2	34,647	0.54 (0.37-0.80)	0.015	0.66 (0.50-0.86)	0.890
Grade 3	16,432	0.59 (0.42-0.83)	0.373	--	
In situ cancers					
All in situ	6,667	0.95 (0.55-1.67)	0.155	--	
Ductal carcinoma in situ	3,510	1.29 (0.51-3.25)	0.019	1.02 (0.44-2.37)	0.096

Abbreviations: CI, confidence interval; ER+/-, estrogen receptor positive/negative; GWAS, genome wide association study; HER2+/-, human epidermal growth factor receptor 2 positive/negative; PR+/-, progesterone receptor positive/negative; SNP, single nucleotide polymorphism.

* SNP rs11012732 was identified by MR-PRESSO as outlying (likely pleiotropic) for analyses of all invasive, ER+, PR+, HR+/HER2-, and low-grade tumors. For analyses of DCIS (for which MR-PRESSO detected pleiotropy, global test p=0.02) SNP rs34517439 (MR-PRESSO $p_{outlier}$=0.08) was identified as likely pleiotropic by inspecting genetic association scatter plots, comparing individual SNP causal effects, and inspecting leave-one-out analyses.

† Causal odds ratios were estimated by inverse-variance weighted Mendelian randomization, using SNPs identified for accelerometer-defined physical activity in a GWAS by Klimentidis et al (5, 6)

‡ p-value associated with the heterogeneity test statistic (Cochran's Q statistic) measuring heterogeneity of causal effects between SNPs

§ vs pre/perimenopausal controls (n=17,686), assigned using age (<50 years) if menopause status was unknown

¶ vs postmenopausal controls (n=36,766), assigned using age (≥50 years) if menopause status was unknown

-- No outlying SNPs were identified.

Table S7. Comparison of results from different Mendelian randomization methods: Association between the secondary instrumental genetic variables for overall physical activity (per standard deviation) and risk of breast cancer

Type of breast cancer	N cases (vs. 54,452 controls)	Odds ratios (95% CI) *†	P for heterogeneity‡ or pleiotropy§
Invasive cancers			
All invasive	69,838		
IVW		0.61 (0.44-0.85)	0.010
Weighted median		0.70 (0.51-0.95)	
MR-Egger		1.28 (0.27-5.98)	0.342
Pre/perimenopausal	¶ 23,999		
IVW		0.65 (0.45-0.93)	0.494
Weighted median		0.77 (0.47-1.29)	
MR-Egger		1.39 (0.24-8.00)	0.378
Postmenopausal	**45,839		
IVW		0.60 (0.43-0.84)	0.068
Weighted median		0.65 (0.45-0.92)	
MR-Egger		1.30 (0.26-6.56)	0.343
By receptor status			
ER+	46,528		
IVW		0.57 (0.39-0.83)	0.004
Weighted median		0.73 (0.53-1.02)	
MR-Egger		0.95 (0.15-6.23)	0.584
ER-	11,246		
IVW		0.72 (0.45-1.17)	0.109
Weighted median		0.63 (0.37-1.08)	
MR-Egger		0.37 (0.03-4.10)	0.575
PR+	34,891		
IVW		0.55 (0.36-0.84)	0.003
Weighted median		0.67 (0.47-0.98)	
MR-Egger		0.83 (0.10-7.07)	0.695
PR-	16,432		
IVW		0.66 (0.48-0.92)	0.443
Weighted median		0.66 (0.42-1.03)	
MR-Egger		0.65 (0.12-3.47)	0.984
HER2+	6,945		
IVW		0.58 (0.34-0.97)	0.254
Weighted median		0.45 (0.23-0.85)	
MR-Egger		0.33 (0.02-4.64)	0.674
HER2-	33,214		
IVW		0.70 (0.50-0.99)	0.072
Weighted median		0.79 (0.54-1.14)	
MR-Egger		1.07 (0.19-5.93)	0.626
Combined hormone receptor- and/or HER2-defined subtypes			
ER+ or PR+; HER2+	4,816		
IVW		0.49 (0.29-0.85)	0.458
Weighted median		0.60 (0.29-1.28)	
MR-Egger		0.08 (0.01-1.16)	0.175
ER+ or PR+; HER2-	27,874		
IVW		0.68 (0.45-1.04)	0.010
Weighted median		0.81 (0.55-1.20)	
MR-Egger		1.15 (0.14-9.73)	0.622

Type of breast cancer	N cases (vs. 54,452 controls)	Odds ratios (95% CI) *†	P for heterogeneity‡ or pleiotropy§
ER-; PR-; HER2+	1,974		
IVW		0.74 (0.33-1.68)	0.468
Weighted median		0.75 (0.25-2.23)	
MR-Egger		2.37 (0.04-129)	0.560
ER-; PR-; HER2-	4,964		
IVW		0.73 (0.36-1.47)	0.088
Weighted median		0.74 (0.35-1.57)	
MR-Egger		0.63 (0.02-21.6)	0.934
ER- and PR- (all)	9,215		
IVW		0.70 (0.42-1.17)	0.126
Weighted median		0.68 (0.38-1.22)	
MR-Egger		0.44 (0.03-5.93)	0.724
By morphology			
Ductal	42,223		
IVW		0.63 (0.46-0.86)	0.067
Weighted median		0.70 (0.50-0.99)	
MR-Egger		0.89 (0.18-4.37)	0.654
Lobular	8,795		
IVW		0.45 (0.28-0.71)	0.358
Weighted median		0.59 (0.32-1.09)	
MR-Egger		1.55 (0.18-13.6)	0.256
By stage at diagnosis			
Stage I	17,583		
IVW		0.59 (0.43-0.82)	0.558
Weighted median		0.63 (0.41-0.99)	
MR-Egger		0.81 (0.17-3.87)	0.694
Stage II	15,992		
IVW		0.52 (0.35-0.78)	0.231
Weighted median		0.51 (0.30-0.87)	
MR-Egger		1.69 (0.26-11.0)	0.206
Stage III/IV	4,553		
IVW		0.51 (0.28-0.93)	0.385
Weighted median		0.41 (0.19-0.89)	
MR-Egger		0.42 (0.02-8.94)	0.891
By tumor grade			
Grade 1/2	34,647		
IVW		0.54 (0.37-0.80)	0.015
Weighted median		0.60 (0.42-0.86)	
MR-Egger		0.73 (0.10-5.21)	0.755
Grade 3	16,432		
IVW		0.59 (0.42-0.83)	0.373
Weighted median		0.55 (0.35-0.88)	
MR-Egger		2.14 (0.45-10.2)	0.099
In situ cancers			
All in situ	6,667		
IVW		0.95 (0.55-1.67)	0.155
Weighted median		0.94 (0.48-1.84)	
MR-Egger		2.38 (0.15-36.5)	0.503
Ductal carcinoma in situ	3,510		
IVW		1.29 (0.51-3.25)	0.019
Weighted median		1.51 (0.59-3.84)	
MR-Egger		0.27 (0.00-27.7)	0.499

Abbreviations: CI, confidence interval; ER+/-, estrogen receptor positive/negative; GWAS, genome wide association study; HER2+/-, human epidermal growth factor receptor 2 positive/negative; IVW, inverse-variance weighted; MR, Mendelian Randomization; PR+/-, progesterone receptor positive/negative; SNP, single nucleotide polymorphism.

* Causal odds ratios were estimated by three different Mendelian randomization methods, using ten SNPs identified for accelerometer-defined physical activity in a GWAS by Klimentidis et al (5, 6)

† Confidence intervals from weighted median MR were generated from 10,000 bootstrap samples (seed 314159265)

‡ Relating to IVW MR estimates: p-value associated with the heterogeneity test statistic (Cochran's Q statistic) measuring heterogeneity of causal effects between SNPs

§ Relating to MR-Egger estimates: p-value for the intercept test (p-value for pleiotropy)

¶ vs pre/perimenopausal controls (n=17,686), assigned using age (<50 years) if menopause status was unknown

** vs postmenopausal controls (n=36,766), assigned using age (≥50 years) if menopause status was unknown

Table S8. Leave-one-out analyses: Association between the secondary instrumental genetic variables for overall physical activity (per standard deviation) and risk of breast cancer, omitting one SNP at a time

	Odds ratios (95% CI) *									
Type of cancer	**Excluding rs12045968**	**Excluding rs34517439 †**	**Excluding rs6775319 §**	**Excluding rs9293503**	**Excluding rs12522261**	**Excluding rs11012732 ¶**	**Excluding rs148193266 ****	**Excluding rs1550435 ††**	**Excluding rs55657917 §§**	**Excluding rs59499656 §**
Invasive cancers										
All invasive	‡0.60 (0.42-0.85)	‡0.58 (0.42-0.81)	‡0.61 (0.43-0.87)	‡0.60 (0.42-0.85)	‡0.61 (0.43-0.87)	0.71 (0.57-0.88)	‡0.59 (0.42-0.83)	‡0.61 (0.42-0.87)	‡0.65 (0.46-0.92)	‡0.59 (0.42-0.85)
Pre/perimenop.	0.68 (0.46-0.99)	0.59 (0.41-0.87)	0.62 (0.42-0.92)	0.61 (0.42-0.90)	0.65 (0.44-0.97)	0.71 (0.49-1.05)	0.62 (0.43-0.91)	0.63 (0.43-0.94)	0.72 (0.49-1.07)	0.62 (0.42-0.92)
Postmenop.	0.57 (0.40-0.82)	0.58 (0.40-0.83)	‡0.60 (0.41-0.88)	‡0.59 (0.41-0.87)	‡0.6 (0.41-0.87)	0.71 (0.54-0.93)	0.58 (0.40-0.83)	‡0.60 (0.41-0.87)	‡0.62 (0.42-0.91)	‡0.59 (0.40-0.85)
By receptor status										
ER+	‡0.56 (0.37-0.85)	‡0.55 (0.37-0.83)	‡0.56 (0.37-0.84)	‡0.55 (0.36-0.83)	‡0.56 (0.37-0.84)	0.69 (0.54-0.88)	‡0.55 (0.37-0.83)	‡0.55 (0.36-0.83)	‡0.60 (0.39-0.91)	‡0.56 (0.37-0.86)
ER-	0.73 (0.43-1.25)	0.65 (0.41-1.03)	0.79 (0.49-1.30)	0.74 (0.43-1.26)	0.77 (0.46-1.29)	0.63 (0.40-1.00)	0.79 (0.49-1.28)	0.73 (0.43-1.25)	0.76 (0.44-1.30)	0.66 (0.40-1.09)
PR+	‡0.53 (0.33-0.86)	‡0.52 (0.33-0.81)	‡0.54 (0.34-0.87)	‡0.54 (0.34-0.87)	‡0.54 (0.33-0.86)	0.67 (0.51-0.88)	‡0.53 (0.33-0.84)	‡0.53 (0.33-0.84)	‡0.59 (0.37-0.94)	‡0.52 (0.33-0.82)
PR-	0.66 (0.46-0.95)	0.63 (0.45-0.89)	0.68 (0.47-0.97)	0.64 (0.45-0.92)	0.74 (0.53-1.05)	0.60 (0.43-0.85)	0.70 (0.49-0.98)	0.67 (0.46-0.96)	0.68 (0.47-0.98)	0.64 (0.45-0.92)
HER2+	0.49 (0.30-0.79)	0.53 (0.31-0.90)	0.55 (0.31-0.97)	0.60 (0.33-1.07)	0.59 (0.33-1.06)	0.62 (0.36-1.10)	0.60 (0.34-1.06)	0.61 (0.35-1.08)	0.68 (0.41-1.13)	0.54 (0.31-0.96)
HER2-	0.72 (0.49-1.05)	0.65 (0.47-0.90)	‡0.70 (0.48-1.03)	0.68 (0.47-1.00)	0.67 (0.46-0.96)	0.81 (0.62-1.07)	‡0.69 (0.48-1.01)	0.68 (0.47-0.99)	0.74 (0.50-1.08)	‡0.69 (0.47-1.02)
Combined hormone receptor- and/or HER2-defined subtypes										
HR+; HER2+	0.42 (0.24-0.75)	0.46 (0.26-0.81)	0.46 (0.26-0.83)	0.52 (0.29-0.95)	0.48 (0.27-0.88)	0.53 (0.29-0.95)	0.53 (0.30-0.94)	0.49 (0.27-0.88)	0.60 (0.34-1.08)	0.47 (0.26-0.87)
HR+; HER2-	‡0.70 (0.43-1.12)	‡0.64 (0.41-0.99)	‡0.66 (0.41-1.05)	‡0.66 (0.41-1.06)	‡0.63 (0.41-0.98)	0.83 (0.62-1.12)	‡0.67 (0.42-1.06)	‡0.67 (0.42-1.07)	‡0.71 (0.44-1.15)	‡0.68 (0.42-1.10)
HR-; HER2+	0.56 (0.24-1.32)	0.67 (0.29-1.58)	0.67 (0.28-1.61)	0.74 (0.30-1.82)	0.88 (0.37-2.09)	0.79 (0.32-1.93)	0.74 (0.30-1.79)	0.88 (0.37-2.09)	0.84 (0.35-2.04)	0.70 (0.29-1.73)
HR-; HER2-	0.76 (0.35-1.64)	0.66 (0.32-1.39)	0.95 (0.54-1.67)	0.73 (0.33-1.60)	0.76 (0.35-1.64)	0.60 (0.31-1.17)	0.73 (0.34-1.57)	0.67 (0.32-1.42)	0.82 (0.38-1.76)	0.69 (0.32-1.50)
HR- (all)	0.69 (0.39-1.22)	0.63 (0.38-1.04)	0.79 (0.48-1.31)	0.71 (0.40-1.26)	0.76 (0.44-1.31)	0.60 (0.37-0.97)	0.74 (0.43-1.28)	0.70 (0.39-1.23)	0.76 (0.44-1.34)	0.65 (0.37-1.12)
By morphology										
Ductal	0.59 (0.43-0.82)	0.60 (0.43-0.84)	‡0.61 (0.43-0.86)	‡0.62 (0.44-0.88)	‡0.62 (0.44-0.88)	0.72 (0.56-0.93)	0.61 (0.43-0.86)	‡0.62 (0.44-0.89)	0.67 (0.48-0.94)	0.60 (0.43-0.85)
Lobular	0.44 (0.26-0.73)	0.41 (0.26-0.66)	0.47 (0.29-0.78)	0.43 (0.26-0.71)	0.51 (0.32-0.81)	0.53 (0.33-0.85)	0.45 (0.27-0.74)	0.42 (0.26-0.68)	0.43 (0.26-0.73)	0.43 (0.26-0.72)
By stage at diagnosis										
Stage I	0.56 (0.40-0.80)	0.59 (0.42-0.83)	0.57 (0.40-0.80)	0.61 (0.43-0.86)	0.61 (0.43-0.87)	0.67 (0.47-0.95)	0.57 (0.40-0.80)	0.58 (0.41-0.82)	0.63 (0.44-0.89)	0.56 (0.40-0.80)
Stage II	0.49 (0.32-0.75)	0.49 (0.32-0.74)	0.54 (0.34-0.84)	0.52 (0.33-0.82)	0.52 (0.33-0.82)	0.59 (0.40-0.86)	0.47 (0.32-0.68)	0.50 (0.32-0.78)	0.58 (0.38-0.88)	0.52 (0.33-0.83)
Stage III/IV	0.42 (0.23-0.77)	0.47 (0.25-0.87)	0.59 (0.32-1.09)	0.52 (0.26-1.01)	0.46 (0.25-0.86)	0.56 (0.29-1.07)	0.53 (0.27-1.02)	0.52 (0.27-1.01)	0.57 (0.30-1.10)	0.53 (0.27-1.04)
By tumor grade										
Grade 1/2	‡0.51 (0.34-0.76)	‡0.54 (0.35-0.83)	‡0.51 (0.34-0.78)	‡0.52 (0.34-0.80)	‡0.54 (0.35-0.83)	0.66 (0.50-0.86)	‡0.53 (0.35-0.82)	‡0.53 (0.34-0.81)	‡0.56 (0.36-0.87)	‡0.54 (0.35-0.83)
Grade 3	0.58 (0.40-0.84)	0.54 (0.38-0.76)	0.63 (0.45-0.90)	0.59 (0.40-0.87)	0.59 (0.4-0.86)	0.64 (0.45-0.91)	0.55 (0.39-0.77)	0.60 (0.41-0.87)	0.62 (0.42-0.90)	0.56 (0.39-0.82)
In situ cancers										
All in situ	0.81 (0.49-1.34)	0.88 (0.49-1.59)	0.91 (0.49-1.67)	0.95 (0.51-1.76)	0.96 (0.51-1.78)	0.97 (0.52-1.81)	0.89 (0.49-1.60)	1.03 (0.57-1.87)	1.17 (0.70-1.94)	1.05 (0.58-1.90)
DCIS	‡1.14 (0.42-3.07)	1.02 (0.44-2.37)	‡1.14 (0.42-3.09)	‡1.24 (0.44-3.50)	‡1.14 (0.42-3.07)	‡1.27 (0.45-3.61)	‡1.47 (0.56-3.84)	‡1.36 (0.49-3.79)	1.79 (0.76-4.23)	‡1.51 (0.56-4.06)

Abbreviations: CI, confidence interval; DCIS, ductal carcinoma in situ; ER+/-, estrogen receptor positive/negative; GWAS, genome wide association study; HER2+/-, human epidermal growth factor receptor 2 positive/negative; PR+/-, progesterone receptor positive/negative; SNP, single nucleotide polymorphism.

* Causal odds ratios were estimated by inverse-variance weighted Mendelian randomization, using SNPs identified for accelerometer-defined physical activity in a GWAS by Klimentidis et al (5, 6)

† This SNP was identified by inspecting scatter plots and individual SNP causal effects as a likely outlier for analyses of DCIS, and was associated in prior GWAS with several possible confounders and with adiposity.

§ This SNP was associated with adiposity in prior GWAS.

¶ This SNP was identified by MR-PRESSO as an outlier for analyses of all invasive, ER+, PR+, HR+/HER2-, and well/moderately differentiated cancers, and was associated in prior GWAS with adiposity and risk of several cancers.

** This SNP had imputation quality score <0.9 and low minor allele frequency (3.1%)

†† This SNP was associated with a possible confounder (height) in prior GWAS.

§§ This SNP was associated with several possible confounders, risk of cancer (ovarian), and with expression of genes associated with multiple relevant traits, including breast cancer risk.

‡ p-value associated with the heterogeneity test statistic (Cochran's Q statistic) measuring heterogeneity of causal effects between SNPs was <0.05

¶

Table S9. Comparison of results from different Mendelian randomization methods: Association between instrumental genetic variables for self-reported vigorous physical activity (≥ 3 vs. 0 days/week) and risk of breast cancer

Type of breast cancer	N cases (vs. 54,452 controls)	Odds ratios (95% CI) *†	P for heterogeneity‡ or pleiotropy§
Invasive cancers			
All invasive	69,838		
IVW		0.83 (0.69-1.01)	0.650
Weighted median		0.80 (0.62-1.02)	
MR-Egger		0.71 (0.18-2.87)	0.821
Pre/perimenopausal	¶ 23,999		
IVW		0.62 (0.45-0.87)	0.788
Weighted median		0.59 (0.39-0.89)	
MR-Egger		0.45 (0.04-5.25)	0.790
Postmenopausal	**45,839		
IVW		0.95 (0.75-1.19)	0.630
Weighted median		0.95 (0.70-1.28)	
MR-Egger		0.95 (0.17-5.28)	0.997
By receptor status			
ER+	46,528		
IVW		0.86 (0.70-1.07)	0.917
Weighted median		0.88 (0.68-1.14)	
MR-Egger		0.90 (0.19-4.25)	0.962
ER-	11,246		
IVW		0.86 (0.61-1.21)	0.418
Weighted median		0.91 (0.58-1.44)	
MR-Egger		0.23 (0.02-2.96)	0.311
PR+	34,891		
IVW		0.77 (0.61-0.98)	0.544
Weighted median		0.81 (0.60-1.09)	
MR-Egger		0.88 (0.16-4.92)	0.886
PR-	16,432		
IVW		0.95 (0.70-1.28)	0.948
Weighted median		0.99 (0.68-1.42)	
MR-Egger		0.59 (0.06-5.35)	0.668
HER2+	6,945		
IVW		0.83 (0.53-1.31)	0.327
Weighted median		0.88 (0.50-1.55)	
MR-Egger		0.04 (0.00-0.88)	0.052
HER2-	33,214		
IVW		0.86 (0.68-1.10)	0.550
Weighted median		0.92 (0.67-1.25)	
MR-Egger		2.10 (0.37-12.1)	0.315
Combined hormone receptor- and/or HER2-defined subtypes			
ER+ or PR+; HER2+	4,816		
IVW		1.00 (0.58-1.70)	0.321
Weighted median		1.18 (0.60-2.31)	
MR-Egger		0.03 (0.00-1.33)	0.069
ER+ or PR+; HER2-	27,874		
IVW		0.82 (0.64-1.06)	0.560
Weighted median		0.87 (0.63-1.21)	
MR-Egger		2.47 (0.39-15.7)	0.241

Type of breast cancer	N cases (vs. 54,452 controls)	Odds ratios (95% CI) *†	P for heterogeneity‡ or pleiotropy§
ER-; PR-; HER2+	1,974		
IVW		0.57 (0.27-1.20)	0.727
Weighted median		0.55 (0.21-1.40)	
MR-Egger		0.05 (0.00-10.3)	0.356
ER-; PR-; HER2-	4,964		
IVW		1.30 (0.79-2.12)	0.593
Weighted median		1.28 (0.68-2.43)	
MR-Egger		1.33 (0.03-50.9)	0.987
ER- and PR- (all)	9,215		
IVW		0.95 (0.66-1.39)	0.559
Weighted median		1.02 (0.63-1.67)	
MR-Egger		0.23 (0.01-3.67)	0.311
By morphology			
Ductal	42,223		
IVW		0.81 (0.65-1.00)	0.932
Weighted median		0.79 (0.61-1.03)	
MR-Egger		0.80 (0.16-3.99)	0.991
Lobular	8,795		
IVW		0.78 (0.53-1.17)	0.809
Weighted median		0.81 (0.49-1.34)	
MR-Egger		0.17 (0.01-3.17)	0.300
By stage at diagnosis			
Stage I	17,583		
IVW		0.88 (0.65-1.19)	0.598
Weighted median		0.78 (0.53-1.15)	
MR-Egger		0.37 (0.04-3.36)	0.435
Stage II	15,992		
IVW		0.82 (0.59-1.14)	0.788
Weighted median		0.79 (0.53-1.19)	
MR-Egger		0.84 (0.08-9.25)	0.991
Stage III/IV	4,553		
IVW		0.75 (0.44-1.27)	0.910
Weighted median		0.85 (0.44-1.62)	
MR-Egger		0.22 (0.00-10.3)	0.528
By tumor grade			
Grade 1/2	34,647		
IVW		0.84 (0.66-1.06)	0.640
Weighted median		0.78 (0.58-1.06)	
MR-Egger		0.41 (0.07-2.32)	0.417
Grade 3	16,432		
IVW		0.99 (0.73-1.33)	0.557
Weighted median		1.13 (0.76-1.70)	
MR-Egger		1.38 (0.15-12.8)	0.767
In situ cancers			
All in situ	6,667		
IVW		0.94 (0.43-2.08)	0.007
Weighted median		1.03 (0.52-2.04)	
MR-Egger		0.39 (0.00-308)	0.795
Ductal carcinoma in situ	3,510		
IVW		0.85 (0.42-1.69)	0.204
Weighted median		0.63 (0.28-1.43)	
MR-Egger		0.06 (0.00-8.63)	0.291

Abbreviations: CI, confidence interval; ER+/-, estrogen receptor positive/negative; GWAS, genome wide association study; HER2+/-, human epidermal growth factor receptor 2 positive/negative; IVW, inverse-variance weighted; MR, Mendelian Randomization; PR+/-, progesterone receptor positive/negative; SNP, single nucleotide polymorphism.

* Causal odds ratios were estimated by three different Mendelian randomization methods, using five SNPs identified in a GWAS of physical activity by Klimentidis et al (5)

† Confidence intervals from weighted median MR were generated from 10,000 bootstrap samples (seed 314159265)

‡ Relating to IVW MR estimates: p-value associated with the heterogeneity test statistic (Cochran's Q statistic) measuring heterogeneity of causal effects between SNPs

§ Relating to MR-Egger estimates: p-value for the intercept test (p-value for pleiotropy)

¶ vs pre/perimenopausal controls (n=17,686), assigned using age (<50 years) if menopause status was unknown

** vs postmenopausal controls (n=36,766), assigned using age (≥50 years) if menopause status was unknown

Table S10. Leave-one-out analyses: Association between instrumental genetic variables for self-reported vigorous physical activity (≥ 3 vs. 0 days/week) and risk of breast cancer, omitting one SNP at a time

	Full instrument (five SNPs)		Excluding rs2764261 *		Excluding rs328902		Excluding rs13243553		Excluding rs3781411		Excluding rs1248860 §	
Type of breast cancer	**Odds ratios (95% CI) †**	**P_{het}‡**	**Odds ratios (95% CI) †**	**P_{het}‡**	**Odds ratios (95% CI) †**	**P_{het}‡**	**Odds ratios (95% CI) †**	**P_{het}‡**	**Odds ratios (95% CI) †**	**P_{het}‡**	**Odds ratios (95% CI) †**	**P_{het}‡**
Invasive cancers												
All invasive	0.83 (0.69-1.01)	0.650	0.86 (0.70-1.06)	0.569	0.80 (0.65-0.99)	0.635	0.79 (0.64-0.98)	0.681	0.84 (0.68-1.04)	0.488	0.87 (0.70-1.08)	0.656
Pre/perimenopausal	0.62 (0.45-0.87)	0.788	0.61 (0.42-0.89)	0.644	0.57 (0.39-0.82)	0.933	0.63 (0.44-0.92)	0.640	0.64 (0.44-0.93)	0.651	0.67 (0.46-0.98)	0.800
Postmenopausal	0.95 (0.75-1.19)	0.630	1.00 (0.77-1.29)	0.607	0.92 (0.71-1.20)	0.490	0.88 (0.68-1.14)	0.809	0.94 (0.73-1.22)	0.461	1.00 (0.77-1.30)	0.596
By receptor status												
ER+	0.86 (0.70-1.07)	0.917	0.87 (0.69-1.11)	0.825	0.86 (0.68-1.08)	0.816	0.83 (0.66-1.05)	0.943	0.86 (0.68-1.09)	0.816	0.90 (0.71-1.14)	0.939
ER-	0.86 (0.61-1.21)	0.418	0.83 (0.54-1.29)	0.282	0.76 (0.52-1.11)	0.607	0.82 (0.54-1.25)	0.305	0.93 (0.63-1.38)	0.399	0.96 (0.65-1.42)	0.488
PR+	0.77 (0.61-0.98)	0.544	0.77 (0.59-0.99)	0.383	0.78 (0.60-1.01)	0.379	0.72 (0.56-0.94)	0.628	0.76 (0.59-0.98)	0.396	0.85 (0.66-1.11)	0.880
PR-	0.95 (0.70-1.28)	0.948	0.94 (0.67-1.31)	0.874	0.91 (0.65-1.27)	0.947	0.93 (0.67-1.30)	0.877	0.98 (0.70-1.37)	0.905	0.99 (0.71-1.39)	0.935
HER2+	0.83 (0.53-1.31)	0.327	0.73 (0.45-1.17)	0.380	0.77 (0.45-1.33)	0.255	0.82 (0.46-1.46)	0.203	1.03 (0.64-1.65)	0.797	0.85 (0.47-1.54)	0.204
HER2-	0.86 (0.68-1.10)	0.550	0.90 (0.69-1.17)	0.460	0.83 (0.64-1.08)	0.473	0.85 (0.65-1.11)	0.396	0.81 (0.62-1.06)	0.574	0.94 (0.72-1.23)	0.723
Combined hormone receptor- and/or HER2-defined subtypes												
ER+/PR+; HER2+	1.00 (0.58-1.70)	0.321	0.84 (0.48-1.45)	0.434	0.94 (0.48-1.83)	0.218	1.05 (0.53-2.06)	0.211	1.26 (0.72-2.21)	0.722	0.94 (0.47-1.88)	0.210
ER+/PR+; HER2-	0.82 (0.64-1.06)	0.560	0.87 (0.66-1.15)	0.540	0.81 (0.61-1.07)	0.404	0.80 (0.60-1.05)	0.438	0.76 (0.57-1.01)	0.679	0.88 (0.67-1.17)	0.602
ER-; PR-; HER2+	0.57 (0.27-1.20)	0.727	0.54 (0.24-1.24)	0.580	0.57 (0.25-1.30)	0.563	0.46 (0.20-1.05)	0.894	0.67 (0.29-1.55)	0.704	0.65 (0.28-1.50)	0.650
ER-; PR-; HER2-	1.30 (0.79-2.12)	0.593	1.23 (0.71-2.12)	0.455	1.11 (0.64-1.91)	0.771	1.37 (0.79-2.37)	0.462	1.29 (0.75-2.24)	0.424	1.52 (0.87-2.64)	0.715
ER- and PR- (all)	0.95 (0.66-1.39)	0.559	0.93 (0.61-1.41)	0.408	0.87 (0.57-1.31)	0.597	0.89 (0.59-1.35)	0.483	1.05 (0.69-1.60)	0.556	1.06 (0.69-1.62)	0.597
By morphology												
Ductal	0.81 (0.65-1.00)	0.932	0.82 (0.64-1.05)	0.861	0.81 (0.64-1.03)	0.840	0.77 (0.61-0.98)	0.987	0.81 (0.63-1.03)	0.840	0.83 (0.65-1.06)	0.896
Lobular	0.78 (0.53-1.17)	0.809	0.78 (0.50-1.22)	0.660	0.74 (0.47-1.15)	0.754	0.79 (0.51-1.24)	0.663	0.88 (0.56-1.39)	0.954	0.74 (0.47-1.16)	0.735
By stage at diagnosis												
Stage I	0.88 (0.65-1.19)	0.598	0.93 (0.66-1.29)	0.501	0.86 (0.62-1.20)	0.449	0.79 (0.57-1.10)	0.908	0.93 (0.67-1.31)	0.525	0.91 (0.65-1.28)	0.459
Stage II	0.82 (0.59-1.14)	0.788	0.75 (0.52-1.08)	0.926	0.83 (0.58-1.19)	0.636	0.84 (0.58-1.20)	0.642	0.81 (0.56-1.18)	0.638	0.89 (0.62-1.29)	0.830
Stage III/IV	0.75 (0.44-1.27)	0.910	0.80 (0.45-1.44)	0.862	0.73 (0.41-1.31)	0.812	0.71 (0.39-1.27)	0.846	0.83 (0.46-1.49)	0.914	0.69 (0.38-1.26)	0.874
By tumor grade												
Grade 1 and 2	0.84 (0.66-1.06)	0.640	0.85 (0.65-1.10)	0.480	0.79 (0.61-1.02)	0.689	0.80 (0.61-1.03)	0.622	0.88 (0.68-1.14)	0.592	0.88 (0.68-1.15)	0.604
Grade 3	0.99 (0.73-1.33)	0.557	0.93 (0.66-1.30)	0.502	1.07 (0.76-1.49)	0.587	0.93 (0.66-1.30)	0.496	0.95 (0.67-1.33)	0.436	1.08 (0.76-1.51)	0.599
In situ cancers												
All in situ	0.94 (0.43-2.08)	0.007	1.30 (0.72-2.34)	0.189	0.76 (0.32-1.78)	0.020	0.93 (0.33-2.56)	0.003	1.05 (0.39-2.83)	0.004	0.77 (0.31-1.93)	0.012
DCIS	0.85 (0.42-1.69)	0.204	0.94 (0.41-2.18)	0.146	0.74 (0.33-1.69)	0.165	0.91 (0.38-2.17)	0.129	1.06 (0.53-2.14)	0.310	0.64 (0.34-1.21)	0.483

Abbreviations: CI, confidence interval; DCIS, ductal carcinoma in situ; ER+/-, estrogen receptor positive/negative; GWAS, genome wide association study; HER2+/-, human epidermal growth factor receptor 2 positive/negative; PR+/-, progesterone receptor positive/negative; SNP, single nucleotide polymorphism.

* This SNP was identified by MR-PRESSO as an outlier for analyses of in situ cancers. This SNP has also been identified in prior GWAS of several possible confounders (age at menarche, height), and of adiposity.

§ This SNP has been associated in prior GWAS with comparative body size (height) at age 10, and past tobacco smoking.

† Causal odds ratios were estimated by inverse-variance weighted Mendelian randomization, using SNPs identified in a GWAS of physical activity by Klimentidis et al (5)

‡ p-value associated with the heterogeneity test statistic (Cochran's Q statistic) measuring heterogeneity of causal effects between SNPs

Table S11. Comparison of results from different Mendelian randomization methods: Association between instrumental genetic variables for sedentary time (per standard deviation) and risk of breast cancer

Type of breast cancer	N cases (vs. 54,452 controls)	Odds ratios (95% CI) *†	P for heterogeneity‡ or pleiotropy§
Invasive cancers			
All invasive	69,838		
IVW		1.20 (0.93-1.55)	0.962
Weighted median		1.23 (0.91-1.67)	
MR-Egger		0.99 (0.20-4.82)	0.806
Pre/perimenopausal	¶ 23,999		
IVW		1.22 (0.78-1.90)	0.589
Weighted median		1.15 (0.66-2.00)	
MR-Egger		1.22 (0.07-20.1)	0.998
Postmenopausal	**45,839		
IVW		1.21 (0.89-1.65)	0.983
Weighted median		1.17 (0.81-1.68)	
MR-Egger		0.75 (0.11-5.24)	0.630
By receptor status			
ER+	46,528		
IVW		1.19 (0.90-1.57)	0.992
Weighted median		1.23 (0.89-1.71)	
MR-Egger		0.75 (0.13-4.38)	0.604
ER-	11,246		
IVW		1.43 (0.90-2.26)	0.926
Weighted median		1.28 (0.73-2.22)	
MR-Egger		1.15 (0.06-21.3)	0.882
PR+	34,891		
IVW		1.19 (0.87-1.63)	0.386
Weighted median		1.29 (0.88-1.91)	
MR-Egger		0.17 (0.02-1.17)	0.046
PR-	16,432		
IVW		1.40 (0.94-2.09)	0.435
Weighted median		1.33 (0.80-2.21)	
MR-Egger		3.85 (0.28-52.7)	0.443
HER2+	6,945		
IVW		1.17 (0.67-2.06)	0.718
Weighted median		1.34 (0.67-2.67)	
MR-Egger		0.24 (0.01-8.33)	0.372
HER2-	33,214		
IVW		1.27 (0.93-1.74)	0.955
Weighted median		1.34 (0.92-1.95)	
MR-Egger		0.57 (0.08-4.15)	0.422
Combined hormone receptor- and/or HER2-defined subtypes			
ER+ or PR+; HER2+	4,816		
IVW		0.86 (0.44-1.67)	0.585
Weighted median		0.78 (0.34-1.79)	
MR-Egger		0.18 (0.00-11.4)	0.452
ER+ or PR+; HER2-	27,874		
IVW		1.12 (0.80-1.56)	0.801
Weighted median		1.12 (0.75-1.68)	
MR-Egger		0.50 (0.06-4.07)	0.444

Type of breast cancer	N cases (vs. 54,452 controls)	Odds ratios (95% CI) *†	P for heterogeneity‡ or pleiotropy§
ER-; PR-; HER2+	1,974		
IVW		1.94 (0.71-5.25)	0.646
Weighted median		1.56 (0.46-5.35)	
MR-Egger		0.06 (0.00-32.0)	0.272
ER-; PR-; HER2-	4,964		
IVW		2.04 (1.06-3.93)	0.500
Weighted median		2.52 (1.10-5.79)	
MR-Egger		0.31 (0.00-19.6)	0.367
ER- and PR- (all)	9,215		
IVW		1.77 (1.07-2.92)	0.819
Weighted median		1.72 (0.93-3.17)	
MR-Egger		1.15 (0.05-27.6)	0.788
By morphology			
Ductal	42,223		
IVW		1.21 (0.91-1.62)	0.992
Weighted median		1.21 (0.86-1.70)	
MR-Egger		1.07 (0.17-6.66)	0.894
Lobular	8,795		
IVW		1.12 (0.66-1.91)	0.695
Weighted median		1.02 (0.53-1.98)	
MR-Egger		0.17 (0.01-4.89)	0.266
By stage at diagnosis			
Stage I	17,583		
IVW		1.62 (0.99-2.65)	0.187
Weighted median		1.33 (0.79-2.23)	
MR-Egger		0.45 (0.02-11.2)	0.428
Stage II	15,992		
IVW		1.23 (0.79-1.90)	0.820
Weighted median		1.36 (0.80-2.31)	
MR-Egger		0.29 (0.02-4.47)	0.294
Stage III/IV	4,553		
IVW		0.91 (0.45-1.84)	0.640
Weighted median		1.02 (0.43-2.43)	
MR-Egger		1.17 (0.01-105)	0.912
By tumor grade			
Grade 1/2	34,647		
IVW		1.15 (0.84-1.57)	0.901
Weighted median		1.15 (0.79-1.67)	
MR-Egger		0.65 (0.09-4.68)	0.568
Grade 3	16,432		
IVW		1.32 (0.88-1.97)	0.967
Weighted median		1.24 (0.77-1.98)	
MR-Egger		0.94 (0.07-11.8)	0.788
In situ cancers			
All in situ	6,667		
IVW		1.75 (1.00-3.07)	0.933
Weighted median		1.79 (0.92-3.50)	
MR-Egger		0.75 (0.02-26.1)	0.637
Ductal carcinoma in situ	3,510		
IVW		2.11 (0.99-4.49)	0.487
Weighted median		2.49 (0.96-6.43)	
MR-Egger		0.23 (0.00-27.0)	0.357

Abbreviations: CI, confidence interval; ER+/-, estrogen receptor positive/negative; GWAS, genome wide association study; HER2+/-, human epidermal growth factor receptor 2 positive/negative; IVW, inverse-variance weighted; MR, Mendelian Randomization; PR+/-, progesterone receptor positive/negative; SNP, single nucleotide polymorphism.

* Causal odds ratios were estimated by three different Mendelian randomization methods, using six SNPs identified in a GWAS of accelerometer-measured movement traits by Doherty et al (3)

† Confidence intervals from weighted median MR were generated from 10,000 bootstrap samples (seed 314159265)

‡ Relating to IVW MR estimates: p-value associated with the heterogeneity test statistic (Cochran's Q statistic) measuring heterogeneity of causal effects between SNPs

§ Relating to MR-Egger estimates: p-value for the intercept test (p-value for pleiotropy)

¶ vs pre/perimenopausal controls (n=17,686), assigned using age (<50 years) if menopause status was unknown

** vs postmenopausal controls (n=36,766), assigned using age (≥50 years) if menopause status was unknown

Table S12. Leave-one-out analyses: Association between instrumental genetic variables for sedentary time (per standard deviation) and risk of breast cancer, omitting one SNP at a time

	Odds ratios (95% CI) * †						
Type of cancer	**Full instrument (six SNPs)**	**Excluding rs61776614 ‡**	**Excluding rs1858242**	**Excluding rs26579 §**	**Excluding rs25981 ****	**Excluding rs6870096 ¶**	**Excluding rs34858520**
Invasive cancers							
All invasive	1.20 (0.93-1.55)	1.22 (0.93-1.60)	1.21 (0.92-1.60)	1.18 (0.89-1.55)	1.26 (0.96-1.67)	1.19 (0.90-1.56)	1.16 (0.88-1.54)
Pre/perimenopausal	1.22 (0.78-1.90)	1.20 (0.74-1.94)	1.30 (0.80-2.11)	1.26 (0.77-2.05)	1.40 (0.86-2.29)	1.10 (0.68-1.78)	1.08 (0.66-1.77)
Postmenopausal	1.21 (0.89-1.65)	1.25 (0.90-1.74)	1.19 (0.84-1.67)	1.15 (0.82-1.62)	1.22 (0.87-1.71)	1.22 (0.87-1.72)	1.22 (0.86-1.71)
By receptor status							
ER+	1.19 (0.90-1.57)	1.21 (0.90-1.64)	1.23 (0.90-1.67)	1.16 (0.85-1.58)	1.17 (0.86-1.60)	1.18 (0.87-1.61)	1.16 (0.85-1.59)
ER-	1.43 (0.90-2.26)	1.47 (0.90-2.42)	1.31 (0.79-2.18)	1.47 (0.89-2.45)	1.53 (0.92-2.55)	1.33 (0.81-2.21)	1.45 (0.87-2.41)
PR+	1.19 (0.87-1.63)	1.33 (0.95-1.85)	1.25 (0.86-1.81)	1.16 (0.79-1.70)	1.17 (0.80-1.73)	1.16 (0.79-1.70)	1.06 (0.76-1.49)
PR-	1.40 (0.94-2.09)	1.30 (0.85-2.01)	1.46 (0.91-2.34)	1.49 (0.94-2.37)	1.33 (0.83-2.12)	1.24 (0.80-1.92)	1.62 (1.04-2.52)
HER2+	1.17 (0.67-2.06)	1.29 (0.70-2.38)	1.18 (0.64-2.20)	1.12 (0.60-2.09)	1.01 (0.54-1.88)	1.34 (0.72-2.48)	1.11 (0.60-2.08)
HER2-	1.27 (0.93-1.74)	1.33 (0.94-1.86)	1.33 (0.94-1.88)	1.25 (0.88-1.77)	1.24 (0.87-1.76)	1.27 (0.90-1.80)	1.21 (0.86-1.72)
Combined hormone receptor- and/or HER2-defined subtypes							
HR+; HER2+	0.86 (0.44-1.67)	0.95 (0.47-1.94)	0.85 (0.41-1.77)	0.74 (0.36-1.53)	0.71 (0.34-1.48)	1.02 (0.49-2.09)	0.93 (0.45-1.93)
HR+; HER2-	1.12 (0.80-1.56)	1.16 (0.81-1.66)	1.20 (0.83-1.73)	1.10 (0.76-1.59)	1.08 (0.74-1.56)	1.15 (0.80-1.65)	1.02 (0.70-1.47)
HR-; HER2+	1.94 (0.71-5.25)	2.38 (0.81-6.95)	2.05 (0.69-6.15)	2.14 (0.71-6.40)	2.01 (0.67-6.05)	1.90 (0.64-5.63)	1.31 (0.43-3.93)
HR-; HER2-	2.04 (1.06-3.93)	2.35 (1.16-4.75)	1.82 (0.88-3.73)	1.99 (0.94-4.20)	1.92 (0.92-4.02)	1.74 (0.85-3.54)	2.55 (1.24-5.26)
HR- (all)	1.77 (1.07-2.92)	1.84 (1.07-3.15)	1.66 (0.96-2.89)	1.83 (1.05-3.17)	1.74 (1.00-3.02)	1.58 (0.91-2.72)	2.00 (1.15-3.48)
By morphology							
Ductal	1.21 (0.91-1.62)	1.22 (0.89-1.66)	1.24 (0.90-1.70)	1.17 (0.85-1.62)	1.25 (0.91-1.72)	1.21 (0.88-1.66)	1.19 (0.86-1.63)
Lobular	1.12 (0.66-1.91)	1.26 (0.71-2.25)	1.11 (0.62-1.99)	0.96 (0.53-1.72)	1.19 (0.66-2.15)	1.04 (0.58-1.87)	1.18 (0.65-2.13)
By stage at diagnosis							
Stage I	1.62 (0.99-2.65)	1.73 (0.98-3.04)	1.76 (0.99-3.12)	1.66 (0.91-3.03)	1.69 (0.93-3.08)	1.70 (0.94-3.05)	1.25 (0.81-1.95)
Stage II	1.23 (0.79-1.90)	1.32 (0.83-2.11)	1.34 (0.83-2.17)	1.16 (0.72-1.87)	1.13 (0.70-1.82)	1.18 (0.73-1.89)	1.25 (0.77-2.02)
Stage III/IV	0.91 (0.45-1.84)	0.90 (0.42-1.92)	0.88 (0.41-1.91)	1.07 (0.49-2.32)	0.81 (0.37-1.77)	0.76 (0.35-1.63)	1.10 (0.51-2.40)
By tumor grade							
Grade 1/2	1.15 (0.84-1.57)	1.18 (0.84-1.65)	1.22 (0.86-1.72)	1.12 (0.79-1.58)	1.19 (0.84-1.69)	1.11 (0.79-1.56)	1.09 (0.77-1.54)
Grade 3	1.32 (0.88-1.97)	1.37 (0.89-2.10)	1.25 (0.80-1.94)	1.25 (0.81-1.95)	1.34 (0.86-2.08)	1.38 (0.89-2.13)	1.35 (0.87-2.10)
In situ cancers							
All in situ	1.75 (1.00-3.07)	1.84 (1.01-3.38)	1.76 (0.95-3.26)	1.74 (0.93-3.22)	1.56 (0.84-2.91)	1.93 (1.04-3.55)	1.69 (0.91-3.15)
DCIS	2.11 (0.99-4.49)	2.55 (1.13-5.73)	1.67 (0.73-3.83)	2.33 (1.00-5.41)	1.91 (0.82-4.45)	2.48 (1.09-5.64)	1.86 (0.81-4.28)

Abbreviations: CI, confidence interval; DCIS, ductal carcinoma in situ; ER+/-, estrogen receptor positive/negative; GWAS, genome wide association study; HER2+/-, human epidermal growth factor receptor 2 positive/negative; PR+/-, progesterone receptor positive/negative; SNP, single nucleotide polymorphism.

* Causal odds ratios were estimated by inverse-variance weighted Mendelian randomization using SNPs identified in a GWAS of accelerometer-measured movement traits by Doherty et al (3)

† All p-values associated with the heterogeneity test statistic (Cochran's Q statistic) measuring heterogeneity of causal effects between SNPs were >0.10

‡ This SNP had imputation quality score <0.9 and low minor allele frequency (6.6%), and was suggested by scatter plots and per-SNP forest plots to be a possible outlier for PR+ analyses.

§ This SNP has been associated with a possible confounder (education) and with adiposity in prior GWAS.

** This is a strand-ambiguous SNP with minor allele frequency near 0.50.

¶ This SNP is correlated with a SNP predicting overall activity, rs6895232 (r^2=0.25 using the National Cancer Institute's LDpair application (2)).

Table S13. Power to detect expected associations* and instrument strength by exposure trait and outcome analysed

	Sample size (% cases)	Power (F-statistic)			
		Overall physical activity instrument †	Vigorous physical activity instrument (accelerometer) ‡	Vigorous physical activity instrument (self-reported) §	Sedentary time instrument **
Type of breast cancer					
Invasive cancers					
All invasive	124,290 (56%)	52% (124.31)	15% (26.05)	37% (81.62)	34% (144.20)
Pre/perimenopausal	41,685 (58%)	22% (42.36)	8% (9.40)	16% (28.04)	14% (49.03)
Postmenopausal	82,605 (55%)	38% (82.96)	12% (17.65)	26% (54.58)	24% (96.17)
By receptor status					
ER+	100,980 (46%)	42% (101.19)	12% (21.35)	30% (66.50)	30% (117.35)
ER-	65,698 (17%)	16% (66.18)	7% (14.24)	12% (43.61)	16% (76.69)
PR+	89,343 (39%)	35% (89.64)	11% (19.00)	25% (58.95)	27% (103.94)
PR-	70,884 (23%)	21% (71.33)	8% (15.28)	15% (46.98)	19% (82.67)
HER2+	61,397 (11%)	12% (61.91)	6% (13.37)	10% (40.82)	12% (71.74)
HER2-	87,666 (38%)	34% (87.98)	11% (18.67)	24% (57.86)	26% (102.01)
Combined hormone receptor- and/or HER2-defined subtypes					
ER+/PR+; HER2+	59,268 (8%)	10% (59.80)	6% (12.94)	8% (39.44)	10% (69.29)
ER+/PR+; HER2-	82,326 (34%)	30% (82.68)	10% (17.59)	22% (54.40)	24% (95.85)
ER-; PR-; HER2+	56,426 (3%)	7% (56.98)	5% (12.37)	6% (37.60)	7% (66.01)
ER-; PR-; HER2-	59,416 (8%)	10% (59.95)	6% (12.97)	8% (39.54)	10% (69.46)
ER- and PR- (all)	63,667 (14%)	14% (64.17)	7% (13.83)	11% (42.30)	14% (74.35)
By morphology					
Ductal	96,675 (44%)	40% (96.92)	12% (20.48)	28% (63.71)	29% (112.9)
Lobular	63,247 (14%)	14% (63.75)	7% (13.75)	11% (42.02)	14% (73.87)
By stage at diagnosis					
Stage I	72,035 (24%)	22% (72.47)	8% (15.52)	16% (47.72)	20% (84.00)
Stage II	70,444 (23%)	21% (70.89)	8% (15.20)	15% (46.69)	19% (82.16)
Stage III/IV	59,005 (8%)	10% (59.54)	6% (12.89)	8% (39.27)	10% (68.98)
By tumor grade					
Grade 1/2	89,099 (39%)	35% (89.40)	11% (18.96)	25% (58.79)	27% (103.66)
Grade 3	70,884 (23%)	21% (71.33)	8% (15.28)	15% (46.98)	19% (82.67)
In situ cancers					
All in situ	61,119 (11%)	12% (61.64)	6% (13.32)	9% (40.64)	12% (71.42)
DCIS	57,962 (6%)	9% (58.51)	6% (12.68)	7% (38.60)	9% (67.78)

* Expected associations (assumed 'true' odds ratio of the outcome variable per standard deviation in the exposure) for these power calculations were odds ratios of 0.70 (for physical activity variables) and 1.30 (for sedentary

behaviour). The alpha level was set at 0.05. Power varied according to the sample size in each analysis (determined by outcome examined) and the proportion of variance explained for the association between each instrument and exposure (R^2_{xz}), detailed below. Power was estimated using the mRnd Mendelian randomization power calculation tool, https://shiny.cnsgenomics.com/mRnd/ (148)

† Calculations based on R^2_{xz} = 0.00099 (0.099% of variance in the exposure explained).

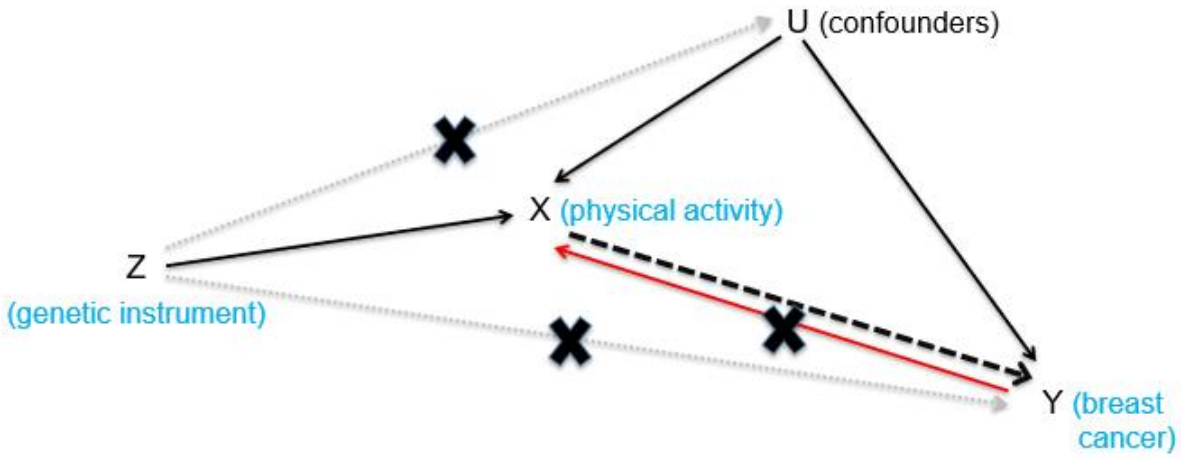
‡ Calculations based on R^2_{xz} = 0.00020 (0.02% of variance in the exposure explained).

§ Calculations based on R^2_{xz} = 0.00065 (0.065% of variance in the exposure explained).

** Calculations based on R^2_{xz} = 0.00115 (0.115% of variance in the exposure explained).

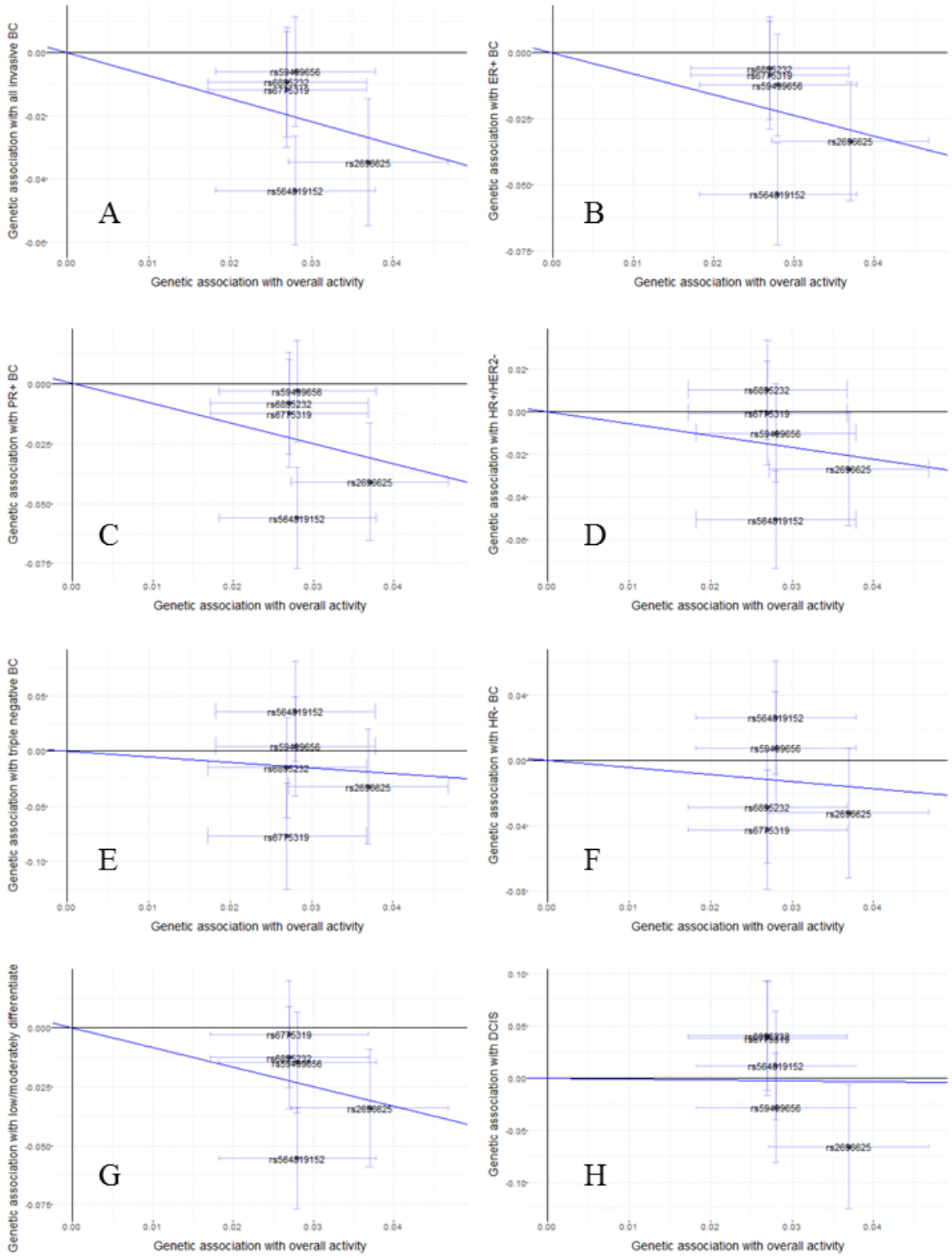
Supplementary Figures

Figure S1. Causal graph of the relationships investigated in this study, illustrating the Mendelian randomization approach and assumptions



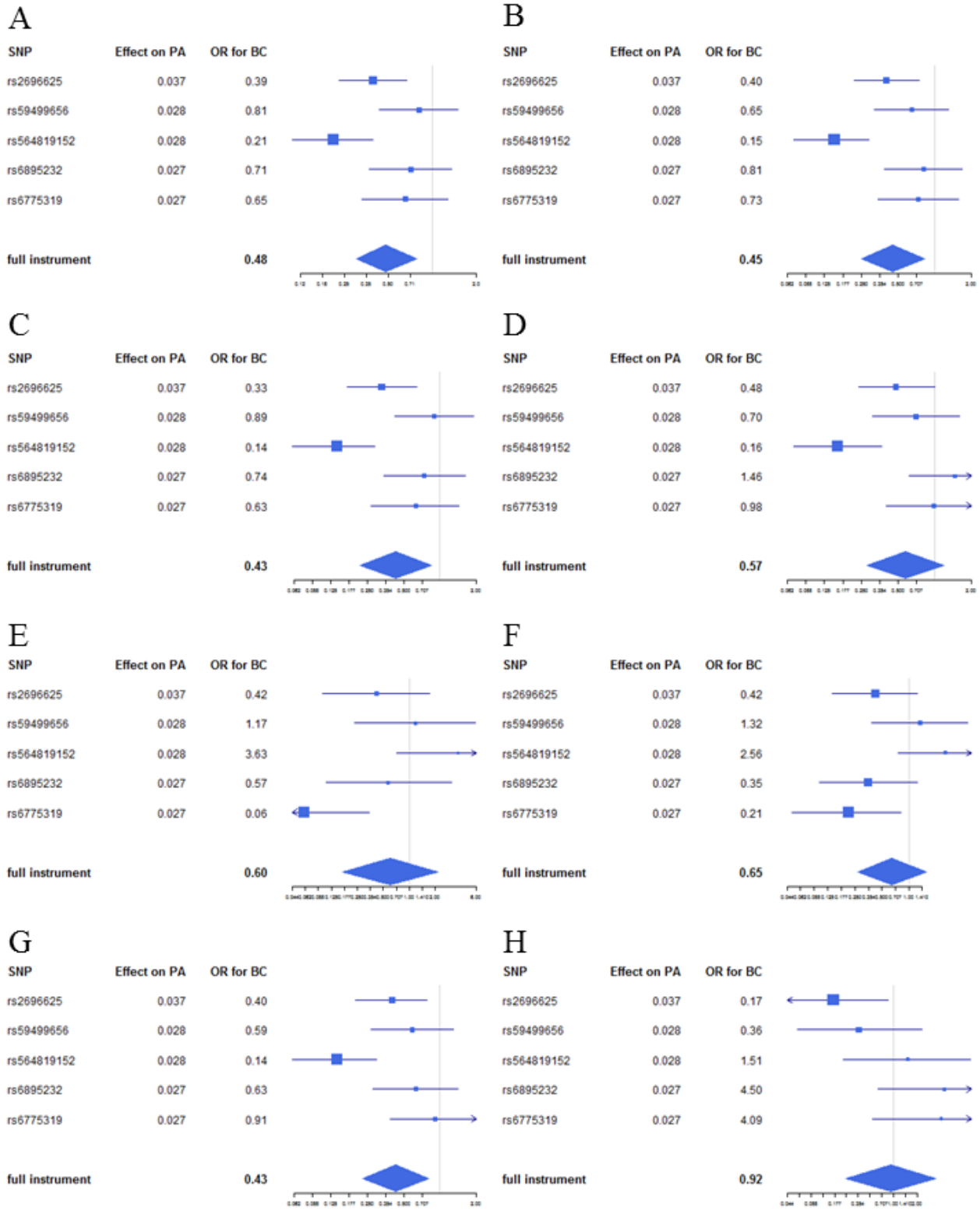
Legend: Crosses indicate an assumption that this causal path does not operate.

Figure S2. Scatter plots of SNP associations with exposure (overall activity; SNPs associated at $p<5x10^{-8}$ (3)) and outcome, for analyses with suspected pleiotropy



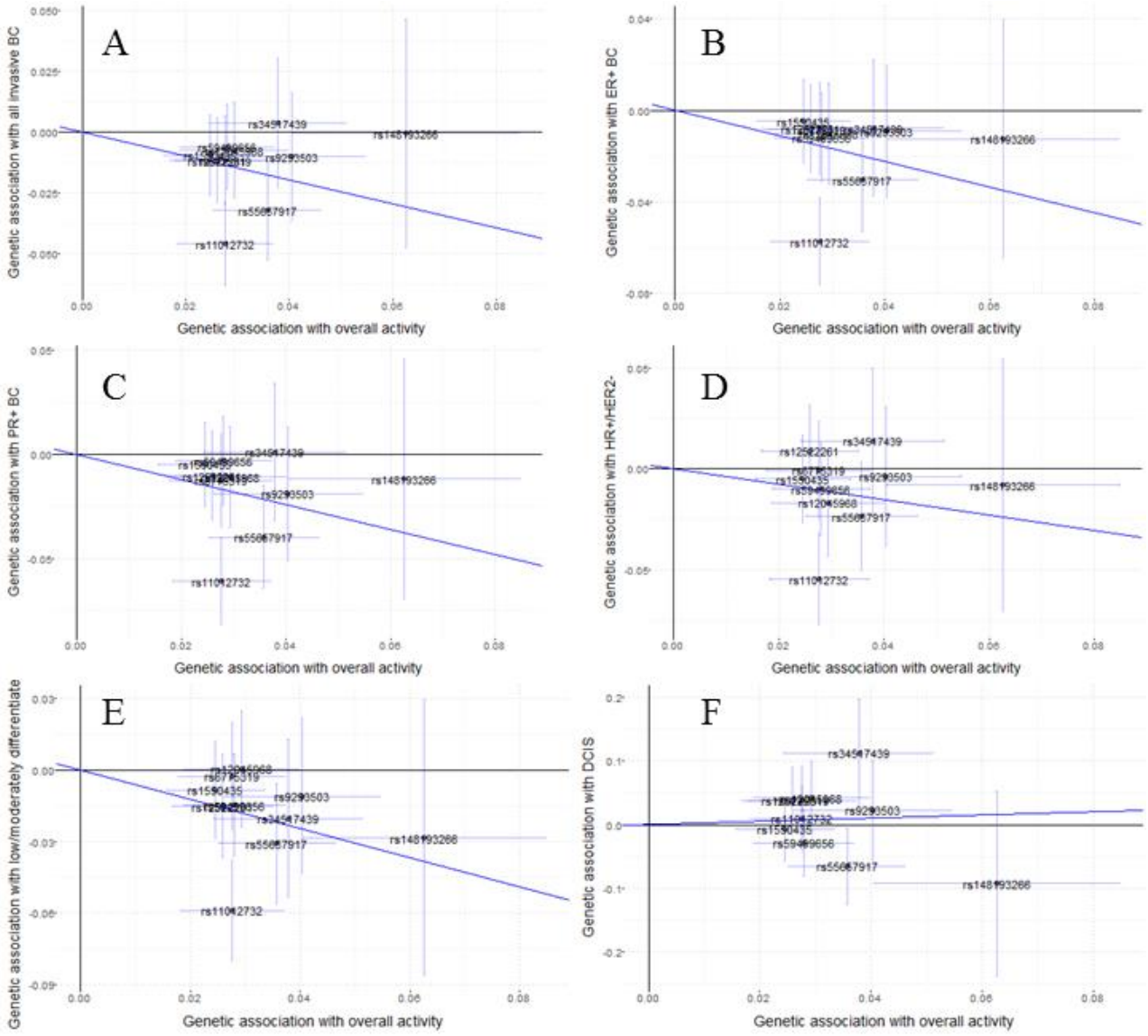
Legend: (A) all invasive breast cancers; (B) ER+; (C) PR+; (D) HR+/HER2-; (E) triple negative; (F) HR-; (G) well/moderately differentiated cancers; (H) DCIS.

Figure S3. Forest plots of individual SNP causal effects on outcomes with suspected pleiotropy: Association between single genetic variants predicting (at $p<5x10^{-8}$) overall physical activity (per standard deviation) and risk of breast cancer



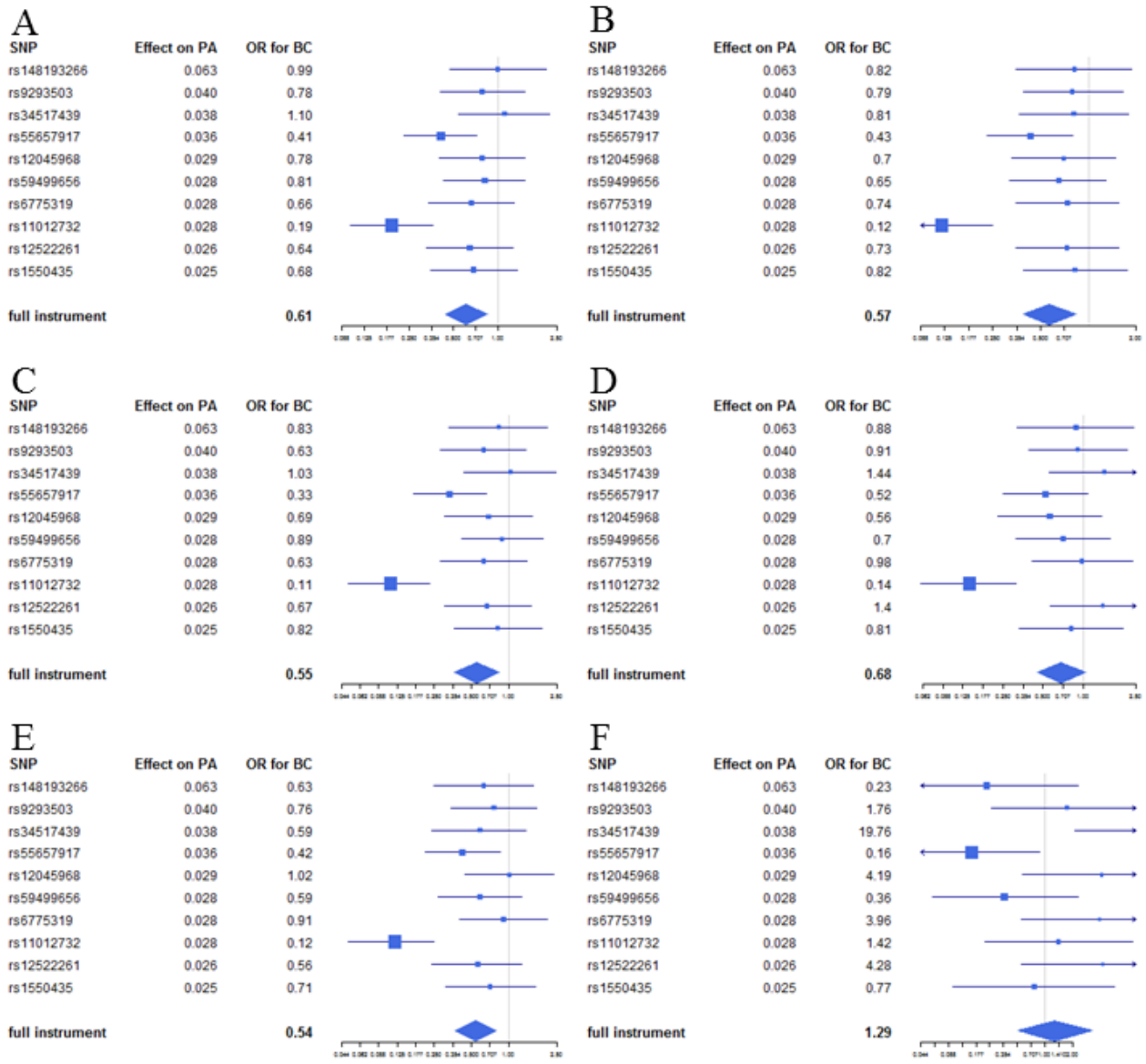
Legend: (A) all invasive breast cancers; (B) ER+; (C) PR+; (D) HR+/HER2-; (E) triple negative; (F) HR-; (G) well/moderately differentiated cancers; (H) DCIS. Associations for each SNP were estimated by Mendelian randomization (Wald ratio technique).

Figure S4. Scatter plots of SNP associations with exposure (overall activity; SNPs associated at $p<5x10^{-7}$ (5, 6)) and outcome, for analyses with suspected pleiotropy



Legend: (A) all invasive breast cancers; (B) ER+; (C) PR+; (D) HR+/HER2-; (E) well/moderately differentiated cancers; (F) DCIS.

Figure S5. Forest plots of individual SNP causal effects on outcomes with suspected pleiotropy: Association between single genetic variants predicting (at $p<5x10^{-7}$) overall physical activity (per standard deviation) and risk of breast cancer



Legend: (A) all invasive breast cancers; (B) ER+; (C) PR+; (D) HR+/HER2-; (E) well/moderately differentiated cancers; (F) DCIS. Associations for each SNP were estimated by Mendelian randomization (Wald ratio technique).

Figure S6. Scatter plot of SNP associations with exposure (self-reported vigorous activity (5)) and in situ cancers (analysis with suspected pleiotropy)

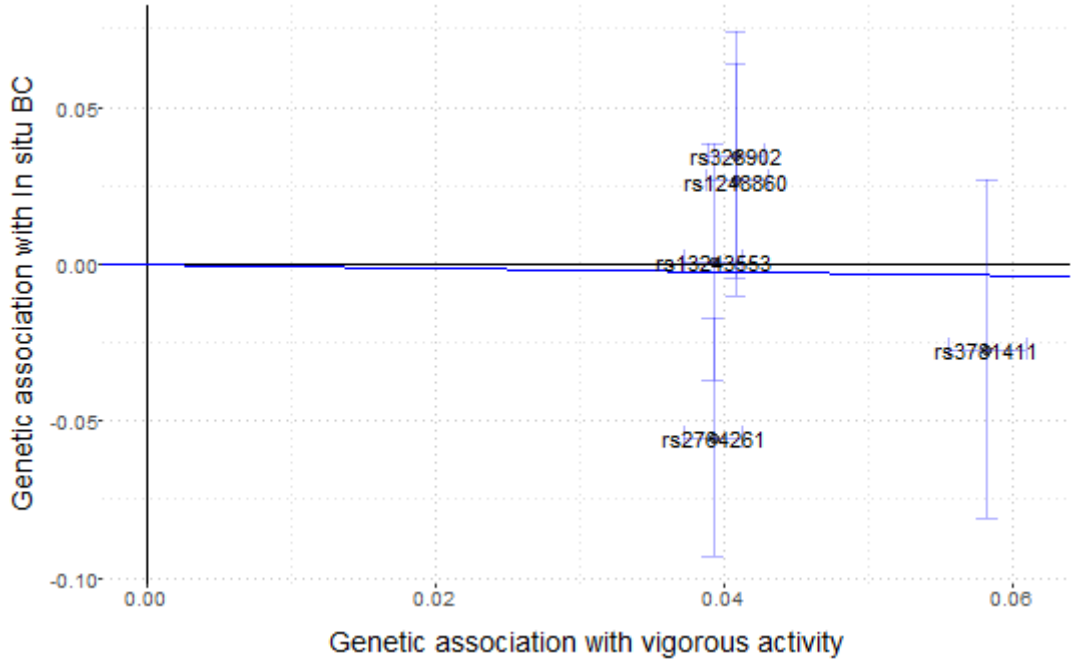
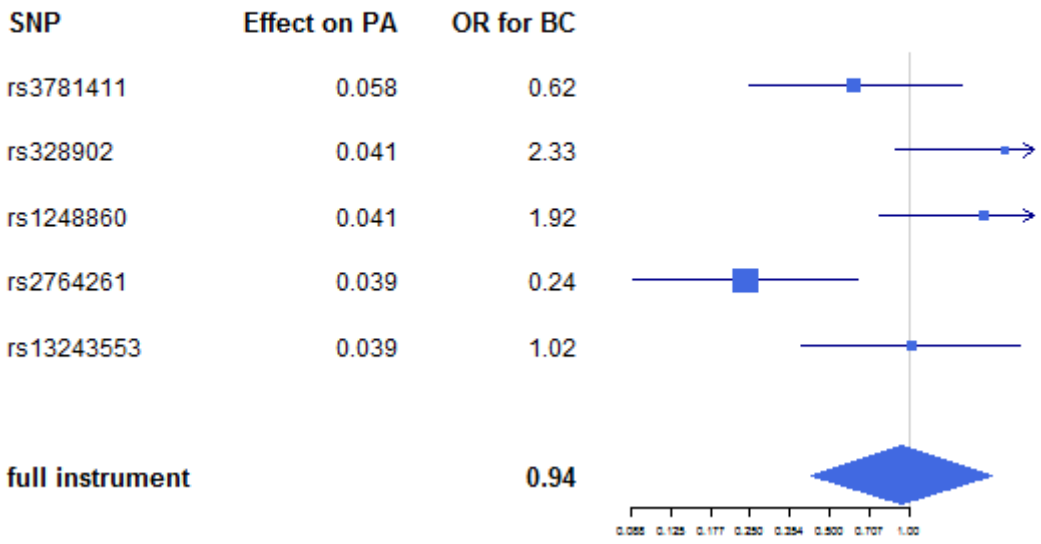


Figure S7. Forest plot of individual SNP causal effects on risk of in situ cancers (suspected pleiotropy): Association between single genetic variants predicting self-reported vigorous physical activity (≥ 3 vs. 0 days/week) and risk of in situ cancers



Note: Associations for each SNP were estimated by Mendelian randomization (Wald ratio technique).

Figure S8. Scatter plot of SNP associations with exposure (sedentary time (3)) and PR+ cancers (analysis with possible pleiotropy)

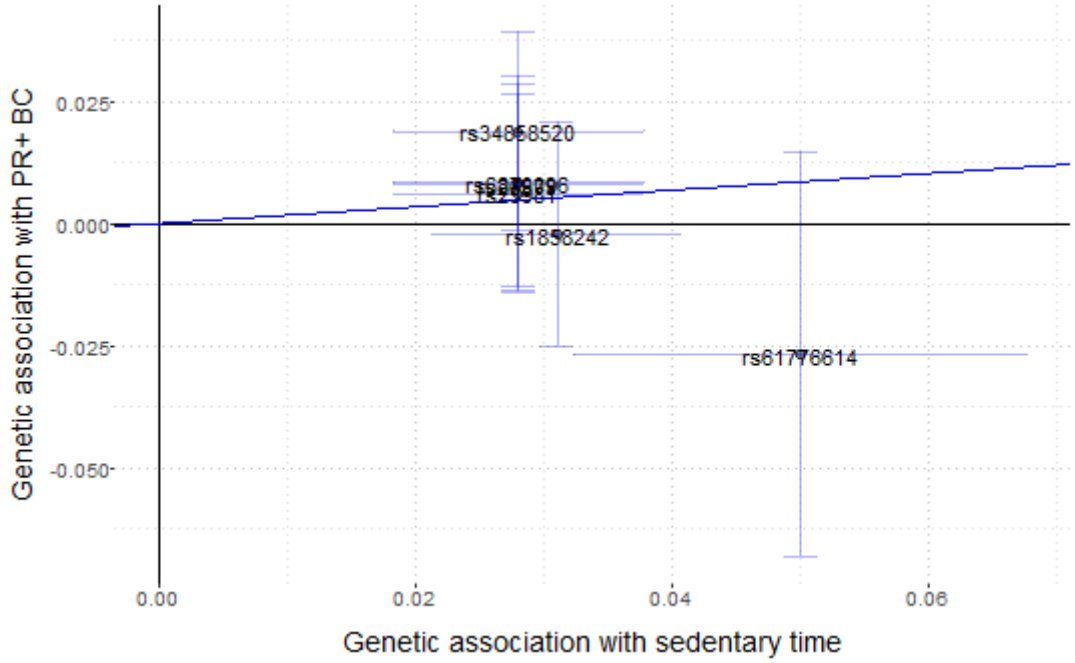
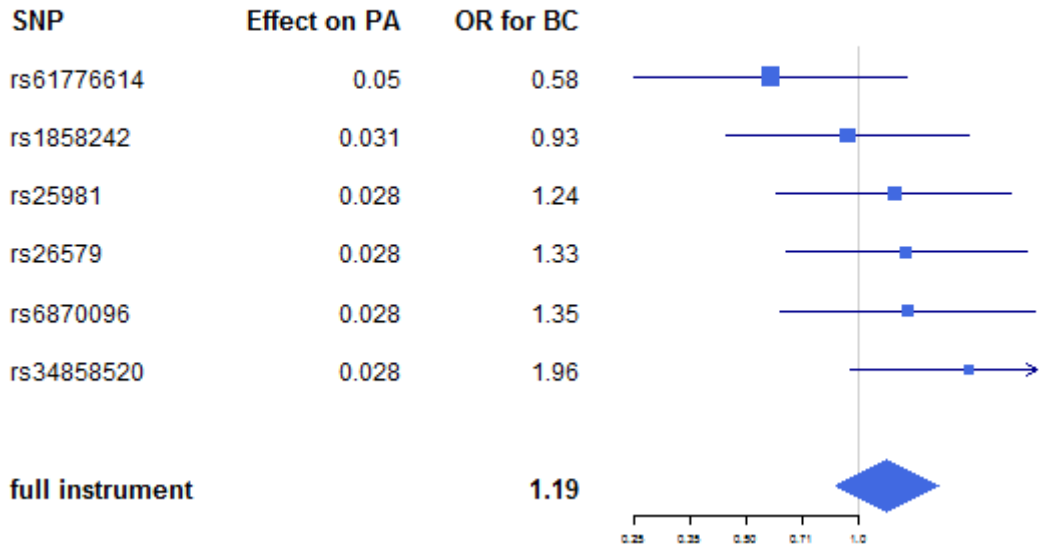


Figure S9. Forest plot of individual SNP causal effects: Association between single genetic variants predicting sedentary time (per standard deviation) and risk of PR+ cancers (possible pleiotropy)



Note: Associations for each SNP were estimated by Mendelian randomization (Wald ratio technique).

Supplementary References

1. Burgess S, Davey Smith G, Davies N, Dudbridge F, Gill D, Glymour MM, et al. Guidelines for performing Mendelian randomization investigations [version 2]. Wellcome Open Res 2020;4:186.
2. Machiela MJ, Chanock SJ. LDlink: A web-based application for exploring population-specific haplotype structure and linking correlated alleles of possible functional variants. Bioinformatics (Oxford, England). 2015;31(21):3555-7.
3. Doherty A, Smith-Byrne K, Ferreira T, Holmes MV, Holmes C, Pulit SL, et al. GWAS identifies 14 loci for device-measured physical activity and sleep duration. Nature communications. 2018;9(1):5257.
4. Doherty A, Jackson D, Hammerla N, Plötz T, Olivier P, Granat MH, et al. Large scale population assessment of physical activity using wrist worn accelerometers: The UK Biobank study. PLoS One. 2017;12(2):e0169649.
5. Klimentidis YC, Raichlen DA, Bea J, Garcia DO, Wineinger NE, Mandarino LJ, et al. Genome-wide association study of habitual physical activity in over 377,000 UK Biobank participants identifies multiple variants including CADM2 and APOE. Int J Obes. 2018;42(6):1161-76.
6. Choi KW, Chen CY, Stein MB, Klimentidis YC, Wang MJ, Koenen KC, et al. Assessment of bidirectional relationships between physical activity and depression among adults: A 2-sample Mendelian randomization study. JAMA psychiatry. 2019;76(4):399-408.
7. Burgess S, Butterworth A, Thompson SG. Mendelian randomization analysis with multiple genetic variants using summarized data. Genetic epidemiology. 2013;37(7):658-65.
8. Ehret GB, Munroe PB, Rice KM, Bochud M, Johnson AD, Chasman DI, et al. Genetic variants in novel pathways influence blood pressure and cardiovascular disease risk. Nature. 2011;478(7367):103-9.
9. Bowden J, Davey Smith G, Haycock PC, Burgess S. Consistent estimation in Mendelian randomization with some invalid instruments using a weighted median estimator. Genetic epidemiology. 2016;40(4):304-14.
10. Bowden J, Davey Smith G, Burgess S. Mendelian randomization with invalid instruments: Effect estimation and bias detection through Egger regression. Int J Epidemiol. 2015;44(2):512-25.
11. Yavorska OO, Burgess S. MendelianRandomization: an R package for performing Mendelian randomization analyses using summarized data. Int J Epidemiol. 2017:1-6.
12. Verbanck M, Chen CY, Neale B, Do R. Detection of widespread horizontal pleiotropy in causal relationships inferred from Mendelian randomization between complex traits and diseases. Nature genetics. 2018;50(5):693-8.
13. Davey Smith G, Davies NM, Dimou N, Egger M, Gallo V, Golub R, et al. STROBE-MR: Guidelines for strengthening the reporting of Mendelian randomization studies. PeerJ Preprints. 2019;7:e27857v1.
14. Fei C, Deroo LA, Sandler DP, Weinberg CR. Fertility drugs and young-onset breast cancer: Results from the Two Sister Study. Journal of the National Cancer Institute. 2012;104(13):1021-7.
15. Dite GS, Jenkins MA, Southey MC, Hocking JS, Giles GG, McCredie MR, et al. Familial risks, early-onset breast cancer, and BRCA1 and BRCA2 germline mutations. Journal of the National Cancer Institute. 2003;95(6):448-57.
16. Schmidt MK, Tollenaar RA, de Kemp SR, Broeks A, Cornelisse CJ, Smit VT, et al. Breast cancer survival and tumor characteristics in premenopausal women carrying the CHEK2*1100delC germline mutation. Journal of clinical oncology : official journal of the American Society of Clinical Oncology. 2007;25(1):64-9.
17. McEvoy M, Smith W, D'Este C, Duke J, Peel R, Schofield P, et al. Cohort profile: The Hunter Community Study. Int J Epidemiol. 2010;39(6):1452-63.
18. Koutros S, Alavanja MC, Lubin JH, Sandler DP, Hoppin JA, Lynch CF, et al. An update of cancer incidence in the Agricultural Health Study. Journal of occupational and environmental medicine. 2010;52(11):1098-105.

19. Lerro CC, Koutros S, Andreotti G, Friesen MC, Alavanja MC, Blair A, et al. Organophosphate insecticide use and cancer incidence among spouses of pesticide applicators in the Agricultural Health Study. Occupational and environmental medicine. 2015;72(10):736-44.
20. Fasching PA, Loehberg CR, Strissel PL, Lux MP, Bani MR, Schrauder M, et al. Single nucleotide polymorphisms of the aromatase gene (CYP19A1), HER2/neu status, and prognosis in breast cancer patients. Breast cancer research and treatment. 2008;112(1):89-98.
21. Schrauder M, Frank S, Strissel PL, Lux MP, Bani MR, Rauh C, et al. Single nucleotide polymorphism D1853N of the ATM gene may alter the risk for breast cancer. Journal of cancer research and clinical oncology. 2008;134(8):873-82.
22. Johnson N, Fletcher O, Naceur-Lombardelli C, dos Santos Silva I, Ashworth A, Peto J. Interaction between CHEK2*1100delC and other low-penetrance breast-cancer susceptibility genes: A familial study. Lancet (London, England). 2005;366(9496):1554-7.
23. Fletcher O, Johnson N, Palles C, dos Santos Silva I, McCormack V, Whittaker J, et al. Inconsistent association between the STK15 F31I genetic polymorphism and breast cancer risk. Journal of the National Cancer Institute. 2006;98(14):1014-8.
24. Fritschi L, Erren TC, Glass DC, Girschik J, Thomson AK, Saunders C, et al. The association between different night shiftwork factors and breast cancer: A case-control study. British journal of cancer. 2013;109(9):2472-80.
25. John EM, Hopper JL, Beck JC, Knight JA, Neuhausen SL, Senie RT, et al. The Breast Cancer Family Registry: An infrastructure for cooperative multinational, interdisciplinary and translational studies of the genetic epidemiology of breast cancer. Breast cancer research : BCR. 2004;6(4):R375-89.
26. Wu HC, Delgado-Cruzata L, Flom JD, Perrin M, Liao Y, Ferris JS, et al. Repetitive element DNA methylation levels in white blood cell DNA from sisters discordant for breast cancer from the New York site of the Breast Cancer Family Registry. Carcinogenesis. 2012;33(10):1946-52.
27. Quante AS, Whittemore AS, Shriver T, Hopper JL, Strauch K, Terry MB. Practical problems with clinical guidelines for breast cancer prevention based on remaining lifetime risk. Journal of the National Cancer Institute. 2015;107(7).
28. Terry MB, Phillips KA, Daly MB, John EM, Andrulis IL, Buys SS, et al. Cohort Profile: The Breast Cancer Prospective Family Study Cohort (ProF-SC). Int J Epidemiol. 2016;45(3):683-92.
29. Rennert G, Pinchev M, Rennert HS. Use of bisphosphonates and risk of postmenopausal breast cancer. Journal of clinical oncology : official journal of the American Society of Clinical Oncology. 2010;28(22):3577-81.
30. Rennert G, Lejbkowicz F, Cohen I, Pinchev M, Rennert HS, Barnett-Griness O. MutYH mutation carriers have increased breast cancer risk. Cancer. 2012;118(8):1989-93.
31. Jiang X, Castelao JE, Chavez-Uribe E, Fernandez Rodriguez B, Celeiro Muñoz C, Redondo CM, et al. Family history and breast cancer hormone receptor status in a Spanish cohort. PLoS One. 2012;7(1):e29459.
32. Redondo CM, Gago-Domínguez M, Ponte SM, Castelo ME, Jiang X, García AA, et al. Breast feeding, parity and breast cancer subtypes in a Spanish cohort. PLoS One. 2012;7(7):e40543.
33. Ali AM, Schmidt MK, Bolla MK, Wang Q, Gago-Dominguez M, Castelao JE, et al. Alcohol consumption and survival after a breast cancer diagnosis: A literature-based meta-analysis and collaborative analysis of data for 29,239 cases. Cancer epidemiology, biomarkers & prevention : a publication of the American Association for Cancer Research, cosponsored by the American Society of Preventive Oncology. 2014;23(6):934-45.
34. Cruz GI, Martínez ME, Natarajan L, Wertheim BC, Gago-Dominguez M, Bondy M, et al. Hypothesized role of pregnancy hormones on HER2+ breast tumor development. Breast cancer research and treatment. 2013;137(1):237-46.
35. Gago-Dominguez M, Castelao JE, Gude F, Fernandez MP, Aguado-Barrera ME, Ponte SM, et al. Alcohol and breast cancer tumor subtypes in a Spanish Cohort. SpringerPlus. 2016;5:39.

36. Yang R, Dick M, Marme F, Schneeweiss A, Langheinz A, Hemminki K, et al. Genetic variants within miR-126 and miR-335 are not associated with breast cancer risk. Breast cancer research and treatment. 2011;127(2):549-54.
37. Grundy A, Schuetz JM, Lai AS, Janoo-Gilani R, Leach S, Burstyn I, et al. Shift work, circadian gene variants and risk of breast cancer. Cancer Epidemiol. 2013;37(5):606-12.
38. Kobayashi LC, Janssen I, Richardson H, Lai AS, Spinelli JJ, Aronson KJ. Moderate-to-vigorous intensity physical activity across the life course and risk of pre- and post-menopausal breast cancer. Breast cancer research and treatment. 2013;139(3):851-61.
39. Grundy A, Richardson H, Burstyn I, Lohrisch C, SenGupta SK, Lai AS, et al. Increased risk of breast cancer associated with long-term shift work in Canada. Occupational and environmental medicine. 2013;70(12):831-8.
40. Kobayashi LC, Janssen I, Richardson H, Lai AS, Spinelli JJ, Aronson KJ. A case-control study of lifetime light intensity physical activity and breast cancer risk. Cancer causes & control : CCC. 2014;25(1):133-40.
41. Menegaux F, Truong T, Anger A, Cordina-Duverger E, Lamkarkach F, Arveux P, et al. Night work and breast cancer: A population-based case-control study in France (the CECILE study). International journal of cancer. 2013;132(4):924-31.
42. Weischer M, Bojesen SE, Tybjaerg-Hansen A, Axelsson CK, Nordestgaard BG. Increased risk of breast cancer associated with CHEK2*1100delC. Journal of clinical oncology : official journal of the American Society of Clinical Oncology. 2007;25(1):57-63.
43. Calle EE, Rodriguez C, Jacobs EJ, Almon ML, Chao A, McCullough ML, et al. The American Cancer Society Cancer Prevention Study II Nutrition Cohort: Rationale, study design, and baseline characteristics. Cancer. 2002;94(9):2490-501.
44. Bernstein L, Allen M, Anton-Culver H, Deapen D, Horn-Ross PL, Peel D, et al. High breast cancer incidence rates among California teachers: Results from the California Teachers Study (United States). Cancer causes & control : CCC. 2002;13(7):625-35.
45. Swann R, Perkins KA, Velentzis LS, Ciria C, Dutton SJ, Mulligan AA, et al. The DietCompLyf study: A prospective cohort study of breast cancer survival and phytoestrogen consumption. Maturitas. 2013;75(3):232-40.
46. Riboli E, Hunt KJ, Slimani N, Ferrari P, Norat T, Fahey M, et al. European Prospective Investigation into Cancer and Nutrition (EPIC): Study populations and data collection. Public health nutrition. 2002;5(6b):1113-24.
47. Widschwendter M, Apostolidou S, Raum E, Rothenbacher D, Fiegl H, Menon U, et al. Epigenotyping in peripheral blood cell DNA and breast cancer risk: A proof of principle study. PLoS One. 2008;3(7):e2656.
48. Evans DG, Astley S, Stavrinos P, Harkness E, Donnelly LS, Dawe S, et al. Improvement in risk prediction, early detection and prevention of breast cancer in the NHS Breast Screening Programme and family history clinics: A dual cohort study. Southampton (UK): NIHR Journals Library; 2016.
49. Ingham SL, Warwick J, Buchan I, Sahin S, O'Hara C, Moran A, et al. Ovarian cancer among 8,005 women from a breast cancer family history clinic: No increased risk of invasive ovarian cancer in families testing negative for BRCA1 and BRCA2. Journal of medical genetics. 2013;50(6):368-72.
50. Kast K, Rhiem K, Wappenschmidt B, Hahnen E, Hauke J, Bluemcke B, et al. Prevalence of BRCA1/2 germline mutations in 21 401 families with breast and ovarian cancer. Journal of medical genetics. 2016;53(7):465-71.
51. Rhiem K, Engel C, Graeser M, Zachariae S, Kast K, Kiechle M, et al. The risk of contralateral breast cancer in patients from BRCA1/2 negative high risk families as compared to patients from BRCA1 or BRCA2 positive families: A retrospective cohort study. Breast cancer research : BCR. 2012;14(6):R156.
52. Graeser MK, Engel C, Rhiem K, Gadzicki D, Bick U, Kast K, et al. Contralateral breast cancer risk in BRCA1 and BRCA2 mutation carriers. Journal of clinical oncology : official journal of the American Society of Clinical Oncology. 2009;27(35):5887-92.

53. Engel C, Rhiem K, Hahnen E, Loibl S, Weber KE, Seiler S, et al. Prevalence of pathogenic BRCA1/2 germline mutations among 802 women with unilateral triple-negative breast cancer without family cancer history. BMC cancer. 2018;18(1):265.
54. Pesch B, Ko Y, Brauch H, Hamann U, Harth V, Rabstein S, et al. Factors modifying the association between hormone-replacement therapy and breast cancer risk. European journal of epidemiology. 2005;20(8):699-711.
55. Justenhoven C, Pierl CB, Haas S, Fischer HP, Baisch C, Hamann U, et al. The CYP1B1_1358_GG genotype is associated with estrogen receptor-negative breast cancer. Breast cancer research and treatment. 2008;111(1):171-7.
56. Denkert C, Von Minckwitz G, Brase JC, Darb-Esfahani S, Gade S, Kronenwett R, et al. Expression of immunologic genes in triple-negative and HER2-positive breast cancer in the neoadjuvant GEPARSIXTO trial: Prediction of response to carboplatin-based chemotherapy. Journal of clinical oncology : official journal of the American Society of Clinical Oncology. 2014;32(15_suppl):510-.
57. von Minckwitz G, Schneeweiss A, Loibl S, Salat C, Denkert C, Rezai M, et al. Neoadjuvant carboplatin in patients with triple-negative and HER2-positive early breast cancer (GeparSixto; GBG 66): A randomised phase 2 trial. The Lancet Oncology. 2014;15(7):747-56.
58. Von Minckwitz G, Schneeweiss A, Salat C, Rezai M, Zahm DM, Klare P, et al. A randomized phase II trial investigating the addition of carboplatin to neoadjuvant therapy for triple-negative and HER2-positive early breast cancer (GeparSixto). Journal of clinical oncology : official journal of the American Society of Clinical Oncology. 2013;31(15_suppl):1004-.
59. Von Minckwitz G, Timms K, Untch M, Elkin EP, Fasching PA, Schneeweiss A, et al. Prediction of pathological complete response (pCR) by Homologous Recombination Deficiency (HRD) after carboplatin-containing neoadjuvant chemotherapy in patients with TNBC: Results from GeparSixto. Journal of clinical oncology : official journal of the American Society of Clinical Oncology. 2015;33(15_suppl):1004-.
60. Chang-Claude J, Eby N, Kiechle M, Bastert G, Becher H. Breastfeeding and breast cancer risk by age 50 among women in Germany. Cancer causes & control : CCC. 2000;11(8):687-95.
61. Dörk T, Bendix R, Bremer M, Rades D, Klöpper K, Nicke M, et al. Spectrum of ATM gene mutations in a hospital-based series of unselected breast cancer patients. Cancer research. 2001;61(20):7608-15.
62. Romero A, García-Sáenz JA, Fuentes-Ferrer M, López Garcia-Asenjo JA, Furió V, Román JM, et al. Correlation between response to neoadjuvant chemotherapy and survival in locally advanced breast cancer patients. Annals of oncology : official journal of the European Society for Medical Oncology. 2013;24(3):655-61.
63. Martin M, Romero A, Cheang MC, López García-Asenjo JA, García-Saenz JA, Oliva B, et al. Genomic predictors of response to doxorubicin versus docetaxel in primary breast cancer. Breast cancer research and treatment. 2011;128(1):127-36.
64. Syrjäkoski K, Vahteristo P, Eerola H, Tamminen A, Kivinummi K, Sarantaus L, et al. Population-based study of BRCA1 and BRCA2 mutations in 1035 unselected Finnish breast cancer patients. Journal of the National Cancer Institute. 2000;92(18):1529-31.
65. Kilpivaara O, Bartkova J, Eerola H, Syrjäkoski K, Vahteristo P, Lukas J, et al. Correlation of CHEK2 protein expression and c.1100delC mutation status with tumor characteristics among unselected breast cancer patients. International journal of cancer. 2005;113(4):575-80.
66. Fagerholm R, Hofstetter B, Tommiska J, Aaltonen K, Vrtel R, Syrjäkoski K, et al. NAD(P)H:quinone oxidoreductase 1 NQO1*2 genotype (P187S) is a strong prognostic and predictive factor in breast cancer. Nature genetics. 2008;40(7):844-53.
67. Bogdanova N, Cybulski C, Bermisheva M, Datsyuk I, Yamini P, Hillemanns P, et al. A nonsense mutation (E1978X) in the ATM gene is associated with breast cancer. Breast cancer research and treatment. 2009;118(1):207-11.

68. Wendt C, Lindblom A, Arver B, von Wachenfeldt A, Margolin S. Tumour spectrum in non-BRCA hereditary breast cancer families in Sweden. Hereditary cancer in clinical practice. 2015;13(1):15.
69. Margolin S, Werelius B, Fornander T, Lindblom A. BRCA1 mutations in a population-based study of breast cancer in Stockholm County. Genetic testing. 2004;8(2):127-32.
70. Gabrielson M, Eriksson M, Hammarström M, Borgquist S, Leifland K, Czene K, et al. Cohort Profile: The Karolinska Mammography Project for Risk Prediction of Breast Cancer (KARMA). Int J Epidemiol. 2017;46(6):1740-1g.
71. Hartikainen JM, Tuhkanen H, Kataja V, Dunning AM, Antoniou A, Smith P, et al. An autosome-wide scan for linkage disequilibrium-based association in sporadic breast cancer cases in eastern Finland: Three candidate regions found. Cancer epidemiology, biomarkers & prevention : a publication of the American Association for Cancer Research, cosponsored by the American Society of Preventive Oncology. 2005;14(1):75-80.
72. Hartikainen JM, Tuhkanen H, Kataja V, Eskelinen M, Uusitupa M, Kosma VM, et al. Refinement of the 22q12-q13 breast cancer-associated region: Evidence of TMPRSS6 as a candidate gene in an eastern Finnish population. Clinical cancer research : an official journal of the American Association for Cancer Research. 2006;12(5):1454-62.
73. Neven P, Brouckaert O, Van Belle V, Vanden Bempt I, Hendrickx W, Cho H, et al. In early-stage breast cancer, the estrogen receptor interacts with correlation between human epidermal growth factor receptor 2 status and age at diagnosis, tumor grade, and lymph node involvement. Journal of clinical oncology : official journal of the American Society of Clinical Oncology. 2008;26(10):1768-9; author reply 9-71.
74. De Maeyer L, Van Limbergen E, De Nys K, Moerman P, Pochet N, Hendrickx W, et al. Does estrogen receptor negative/progesterone receptor positive breast carcinoma exist? Journal of clinical oncology : official journal of the American Society of Clinical Oncology. 2008;26(2):335-6; author reply 6-8.
75. Flesch-Janys D, Slanger T, Mutschelknauss E, Kropp S, Obi N, Vettorazzi E, et al. Risk of different histological types of postmenopausal breast cancer by type and regimen of menopausal hormone therapy. International journal of cancer. 2008;123(4):933-41.
76. De Vecchi G, Verderio P, Pizzamiglio S, Manoukian S, Barile M, Fortuzzi S, et al. Evidences for association of the CASP8 -652 6N del promoter polymorphism with age at diagnosis in familial breast cancer cases. Breast cancer research and treatment. 2009;113(3):607-8.
77. Catucci I, Verderio P, Pizzamiglio S, Manoukian S, Peissel B, Barile M, et al. SNPs in ultraconserved elements and familial breast cancer risk. Carcinogenesis. 2009;30(3):544-5; author reply 6.
78. Olson JE, Ma CX, Pelleymounter LL, Schaid DJ, Pankratz VS, Vierkant RA, et al. A comprehensive examination of CYP19 variation and breast density. Cancer epidemiology, biomarkers & prevention : a publication of the American Association for Cancer Research, cosponsored by the American Society of Preventive Oncology. 2007;16(3):623-5.
79. Giles GG, English DR. The Melbourne Collaborative Cohort Study. IARC scientific publications. 2002;156:69-70.
80. Kolonel LN, Henderson BE, Hankin JH, Nomura AM, Wilkens LR, Pike MC, et al. A multiethnic cohort in Hawaii and Los Angeles: Baseline characteristics. American journal of epidemiology. 2000;151(4):346-57.
81. Olsson HL, Ingvar C, Bladström A. Hormone replacement therapy containing progestins and given continuously increases breast carcinoma risk in Sweden. Cancer. 2003;97(6):1387-92.
82. Nielsen K, Måsbäck A, Olsson H, Ingvar C. A prospective, population-based study of 40,000 women regarding host factors, UV exposure and sunbed use in relation to risk and anatomic site of cutaneous melanoma. International journal of cancer. 2012;131(3):706-15.
83. Olson JE, Sellers TA, Scott CG, Schueler BA, Brandt KR, Serie DJ, et al. The influence of mammogram acquisition on the mammographic density and breast cancer association in the Mayo Mammography Health Study cohort. Breast cancer research : BCR. 2012;14(6):R147.

84. Comen E, Balistreri L, Gönen M, Dutra-Clarke A, Fazio M, Vijai J, et al. Discriminatory accuracy and potential clinical utility of genomic profiling for breast cancer risk in BRCA-negative women. Breast cancer research and treatment. 2011;127(2):479-87.
85. Aure MR, Jernström S, Krohn M, Vollan HK, Due EU, Rødland E, et al. Integrated analysis reveals microRNA networks coordinately expressed with key proteins in breast cancer. Genome medicine. 2015;7(1):21.
86. Fleischer T, Frigessi A, Johnson KC, Edvardsen H, Touleimat N, Klajic J, et al. Genome-wide DNA methylation profiles in progression to in situ and invasive carcinoma of the breast with impact on gene transcription and prognosis. Genome biology. 2014;15(8):435.
87. Fleischer T, Edvardsen H, Solvang HK, Daviaud C, Naume B, Børresen-Dale A-L, et al. Integrated analysis of high-resolution DNA methylation profiles, gene expression, germline genotypes and clinical end points in breast cancer patients. 2014;134(11):2615-25.
88. Quigley DA, Fiorito E, Nord S, Van Loo P, Alnæs GG, Fleischer T, et al. The 5p12 breast cancer susceptibility locus affects MRPS30 expression in estrogen-receptor positive tumors. Molecular oncology. 2014;8(2):273-84.
89. Zheng W, Long J, Gao YT, Li C, Zheng Y, Xiang YB, et al. Genome-wide association study identifies a new breast cancer susceptibility locus at 6q25.1. Nature genetics. 2009;41(3):324-8.
90. Newman B, Moorman PG, Millikan R, Qaqish BF, Geradts J, Aldrich TE, et al. The Carolina Breast Cancer Study: Integrating population-based epidemiology and molecular biology. Breast cancer research and treatment. 1995;35(1):51-60.
91. Millikan R, Eaton A, Worley K, Biscocho L, Hodgson E, Huang WY, et al. HER2 codon 655 polymorphism and risk of breast cancer in African Americans and whites. Breast cancer research and treatment. 2003;79(3):355-64.
92. Hankinson SE, Willett WC, Manson JE, Colditz GA, Hunter DJ, Spiegelman D, et al. Plasma sex steroid hormone levels and risk of breast cancer in postmenopausal women. Journal of the National Cancer Institute. 1998;90(17):1292-9.
93. Nurses' Health Study 2020 [Available from: http://www.channing.harvard.edu/nhs/.
94. Tworoger SS, Missmer SA, Eliassen AH, Spiegelman D, Folkerd E, Dowsett M, et al. The association of plasma DHEA and DHEA sulfate with breast cancer risk in predominantly premenopausal women. Cancer epidemiology, biomarkers & prevention : a publication of the American Association for Cancer Research, cosponsored by the American Society of Preventive Oncology. 2006;15(5):967-71.
95. de Bock GH, Schutte M, Krol-Warmerdam EM, Seynaeve C, Blom J, Brekelmans CT, et al. Tumour characteristics and prognosis of breast cancer patients carrying the germline CHEK2*1100delC variant. Journal of medical genetics. 2004;41(10):731-5.
96. Huijts PE, Vreeswijk MP, Kroeze-Jansema KH, Jacobi CE, Seynaeve C, Krol-Warmerdam EM, et al. Clinical correlates of low-risk variants in FGFR2, TNRC9, MAP3K1, LSP1 and 8q24 in a Dutch cohort of incident breast cancer cases. Breast cancer research : BCR. 2007;9(6):R78.
97. García-Closas M, Egan KM, Newcomb PA, Brinton LA, Titus-Ernstoff L, Chanock S, et al. Polymorphisms in DNA double-strand break repair genes and risk of breast cancer: two population-based studies in USA and Poland, and meta-analyses. Human genetics. 2006;119(4):376-88.
98. Pfeiffer RM, Park Y, Kreimer AR, Lacey JV, Jr., Pee D, Greenlee RT, et al. Risk prediction for breast, endometrial, and ovarian cancer in white women aged 50 y or older: Derivation and validation from population-based cohort studies. PLoS medicine. 2013;10(7):e1001492.
99. Eccles D, Gerty S, Simmonds P, Hammond V, Ennis S, Altman DG. Prospective study of outcomes in sporadic versus hereditary breast cancer (POSH): Study protocol. BMC cancer. 2007;7:160.
100. Tapper W, Hammond V, Gerty S, Ennis S, Simmonds P, Collins A, et al. The influence of genetic variation in 30 selected genes on the clinical characteristics of early onset breast cancer. Breast cancer research : BCR. 2008;10(6):R108.

101. Copson E, Eccles B, Maishman T, Gerty S, Stanton L, Cutress RI, et al. Prospective observational study of breast cancer treatment outcomes for UK women aged 18-40 years at diagnosis: the POSH study. Journal of the National Cancer Institute. 2013;105(13):978-88.
102. Copson E, Maishman T, Gerty S, Eccles B, Stanton L, Cutress RI, et al. Ethnicity and outcome of young breast cancer patients in the United Kingdom: the POSH study. British journal of cancer. 2014;110(1):230-41.
103. Copson ER, Cutress RI, Maishman T, Eccles BK, Gerty S, Stanton L, et al. Obesity and the outcome of young breast cancer patients in the UK: the POSH study. Annals of oncology : official journal of the European Society for Medical Oncology. 2015;26(1):101-12.
104. Couch FJ, Hart SN, Sharma P, Toland AE, Wang X, Miron P, et al. Inherited mutations in 17 breast cancer susceptibility genes among a large triple-negative breast cancer cohort unselected for family history of breast cancer. Journal of clinical oncology : official journal of the American Society of Clinical Oncology. 2015;33(4):304-11.
105. Wunderle M, Pretscher J, Brucker SY, Volz B, Hartmann A, Fiessler C, et al. Association between breast cancer risk factors and molecular type in postmenopausal patients with hormone receptor-positive early breast cancer. Breast cancer research and treatment. 2019;174(2):453-61.
106. Kriege M, Hollestelle A, Jager A, Huijts PE, Berns EM, Sieuwerts AM, et al. Survival and contralateral breast cancer in CHEK2 1100delC breast cancer patients: Impact of adjuvant chemotherapy. British journal of cancer. 2014;111(5):1004-13.
107. MacPherson G, Healey CS, Teare MD, Balasubramanian SP, Reed MW, Pharoah PD, et al. Association of a common variant of the CASP8 gene with reduced risk of breast cancer. Journal of the National Cancer Institute. 2004;96(24):1866-9.
108. Rafii S, O'Regan P, Xinarianos G, Azmy I, Stephenson T, Reed M, et al. A potential role for the XRCC2 R188H polymorphic site in DNA-damage repair and breast cancer. Human molecular genetics. 2002;11(12):1433-8.
109. Lesueur F, Pharoah PD, Laing S, Ahmed S, Jordan C, Smith PL, et al. Allelic association of the human homologue of the mouse modifier Ptprj with breast cancer. Human molecular genetics. 2005;14(16):2349-56.
110. Nichols HB, Baird DD, DeRoo LA, Kissling GE, Sandler DP. Tubal ligation in relation to menopausal symptoms and breast cancer risk. British journal of cancer. 2013;109(5):1291-5.
111. Xu Z, Bolick SC, DeRoo LA, Weinberg CR, Sandler DP, Taylor JA. Epigenome-wide association study of breast cancer using prospectively collected sister study samples. Journal of the National Cancer Institute. 2013;105(10):694-700.
112. The Sister Study 2020 [Available from: http://sisterstudy.niehs.nih.gov/.
113. Stevens KN, Fredericksen Z, Vachon CM, Wang X, Margolin S, Lindblom A, et al. 19p13.1 is a triple-negative-specific breast cancer susceptibility locus. Cancer research. 2012;72(7):1795-803.
114. Suzuki R, Ye W, Rylander-Rudqvist T, Saji S, Colditz GA, Wolk A. Alcohol and postmenopausal breast cancer risk defined by estrogen and progesterone receptor status: A prospective cohort study. Journal of the National Cancer Institute. 2005;97(21):1601-8.
115. Andergassen U, Neugebauer J, Janni W, Hepp P, Ortmann U, Sommer H, et al. P276 Simultaneous study of gemcitabine-docetaxel combination adjuvant treatment, as well as biological targeted treatment: the SUCCESS B Trial. The Breast. 2011;20:S66.
116. Jaeger BAS, Rack BK, Neugebauer JK, Andergassen U, Melcher CA, Mouarrawy D, et al. HER2 expression on circulating tumor cells (CTC) in patients with early HER2-positive breast cancer: Results of the German SUCCESS B trial. 2012;30(15_suppl):10540-.
117. Neugebauer JK, Rack BK, Jaeger BAS, Andergassen U, Pestka A, Fasching PA, et al. Persistence of HER2 overexpression on circulating tumor cells in patients after systemic treatment for HER2-positive breast cancer: Follow-up results of the German Success B trial. Journal of clinical oncology : official journal of the American Society of Clinical Oncology. 2013;31(15_suppl):11043-.
118. Hepp P, Sommer H, Beckmann MW, Lichtenegger W, Hauner H, Schneeweiss A, et al. Simultaneous study of docetaxel based anthracycline free adjuvant treatment evaluation, as well as

lifestyle intervention strategies Success C-Studie (1st Capital Congress of the Gynecologic Oncology of the North-Eastern German Society of Gynecological Oncology). Anticancer research. 2009;29(5):1567-8.
119. Jaeger BA, Andergassen U, Neugebauer JK, Alunni-Fabbroni M, Melcher CA, Hagenbeck C, et al. Abstract P4-01-08: Persistence of circulating tumor cells immediately after and two years after systemic adjuvant chemotherapy in patients with early breast cancer – Results of the German SUCCESS trials. 2015;75(9 Supplement):P4-01-8-P4--8.
120. Ortmann U, Salmen J, Hepp PGM, Beckmann MW, Fehm TN, Hindenburg H, et al. The SUCCESS-C trial: Interim analysis of toxicity evaluating the role of an anthracycline-free chemotherapy regimen in the adjuvant treatment of HER2/neu-negative breast cancer. Journal of clinical oncology : official journal of the American Society of Clinical Oncology. 2011;29(15_suppl):1070-.
121. Rack B, Andergassen U, Neugebauer J, Salmen J, Hepp P, Sommer H, et al. The German SUCCESS C Study - The first European lifestyle study on breast cancer. Breast Care. 2010;5(6):395-400.
122. Jakubowska A, Cybulski C, Szymańska A, Huzarski T, Byrski T, Gronwald J, et al. BARD1 and breast cancer in Poland. Breast cancer research and treatment. 2008;107(1):119-22.
123. Jakubowska A, Jaworska K, Cybulski C, Janicka A, Szymańska-Pasternak J, Lener M, et al. Do BRCA1 modifiers also affect the risk of breast cancer in non-carriers? European journal of cancer (Oxford, England : 1990). 2009;45(5):837-42.
124. Cybulski C, Kluźniak W, Huzarski T, Wokołorczyk D, Kashyap A, Jakubowska A, et al. Clinical outcomes in women with breast cancer and a PALB2 mutation: A prospective cohort analysis. The Lancet Oncology. 2015;16(6):638-44.
125. Cybulski C, Carrot-Zhang J, Kluźniak W, Rivera B, Kashyap A, Wokołorczyk D, et al. Germline RECQL mutations are associated with breast cancer susceptibility. Nature genetics. 2015;47(6):643-6.
126. Madsen MJ, Knight S, Sweeney C, Factor R, Salama M, Stijleman IJ, et al. Reparameterization of PAM50 expression identifies novel breast tumor dimensions and leads to discovery of a genome-wide significant breast cancer locus at 12q15. Cancer epidemiology, biomarkers & prevention : a publication of the American Association for Cancer Research, cosponsored by the American Society of Preventive Oncology. 2018;27(6):644-52.
127. Camp NJ, Parry M, Knight S, Abo R, Elliott G, Rigas SH, et al. Fine-mapping CASP8 risk variants in breast cancer. Cancer epidemiology, biomarkers & prevention : a publication of the American Association for Cancer Research, cosponsored by the American Society of Preventive Oncology. 2012;21(1):176-81.
128. Anton-Culver H, Cohen PF, Gildea ME, Ziogas A. Characteristics of BRCA1 mutations in a population-based case series of breast and ovarian cancer. European journal of cancer (Oxford, England : 1990). 2000;36(10):1200-8.
129. Ziogas A, Gildea M, Cohen P, Bringman D, Taylor TH, Seminara D, et al. Cancer risk estimates for family members of a population-based family registry for breast and ovarian cancer. Cancer epidemiology, biomarkers & prevention : a publication of the American Association for Cancer Research, cosponsored by the American Society of Preventive Oncology. 2000;9(1):103-11.
130. Swerdlow AJ, Jones ME, Schoemaker MJ, Hemming J, Thomas D, Williamson J, et al. The Breakthrough Generations Study: Design of a long-term UK cohort study to investigate breast cancer aetiology. British journal of cancer. 2011;105(7):911-7.
131. Balogun N, Gentry-Maharaj A, Wozniak EL, Lim A, Ryan A, Ramus SJ, et al. Recruitment of newly diagnosed ovarian cancer patients proved challenging in a multicentre biobanking study. Journal of clinical epidemiology. 2011;64(5):525-30.
132. Doody MM, Freedman DM, Alexander BH, Hauptmann M, Miller JS, Rao RS, et al. Breast cancer incidence in U.S. radiologic technologists. Cancer. 2006;106(12):2707-15.
133. Bhatti P, Doody MM, Alexander BH, Yuenger J, Simon SL, Weinstock RM, et al. Breast cancer risk polymorphisms and interaction with ionizing radiation among U.S. radiologic technologists. Cancer epidemiology, biomarkers & prevention : a publication of the American Association for Cancer Research, cosponsored by the American Society of Preventive Oncology. 2008;17(8):2007-11.

134. Rajaraman P, Bhatti P, Doody MM, Simon SL, Weinstock RM, Linet MS, et al. Nucleotide excision repair polymorphisms may modify ionizing radiation-related breast cancer risk in US radiologic technologists. International journal of cancer. 2008;123(11):2713-6.
135. Sigurdson AJ, Bhatti P, Chang SC, Rajaraman P, Doody MM, Bowen L, et al. Polymorphisms in estrogen biosynthesis and metabolism-related genes, ionizing radiation exposure, and risk of breast cancer among US radiologic technologists. Breast cancer research and treatment. 2009;118(1):177-84.
136. Neale B. UK Biobank GWAS round 2 2018 [Available from: http://www.nealelab.is/uk-biobank/.
137. Phelan CM, Kuchenbaecker KB, Tyrer JP, Kar SP, Lawrenson K, Winham SJ, et al. Identification of 12 new susceptibility loci for different histotypes of epithelial ovarian cancer. Nature genetics. 2017;49(5):680-91.
138. Tsoi LC, Stuart PE, Tian C, Gudjonsson JE, Das S, Zawistowski M, et al. Large scale meta-analysis characterizes genetic architecture for common psoriasis associated variants. Nature communications. 2017;8:15382.
139. McKay JD, Hung RJ, Han Y, Zong X, Carreras-Torres R, Christiani DC, et al. Large-scale association analysis identifies new lung cancer susceptibility loci and heterogeneity in genetic susceptibility across histological subtypes. Nature genetics. 2017;49(7):1126-32.
140. Dobbins SE, Broderick P, Melin B, Feychting M, Johansen C, Andersson U, et al. Common variation at 10p12.31 near MLLT10 influences meningioma risk. Nature genetics. 2011;43(9):825-7.
141. Wood AR, Esko T, Yang J, Vedantam S, Pers TH, Gustafsson S, et al. Defining the role of common variation in the genomic and biological architecture of adult human height. Nature genetics. 2014;46(11):1173-86.
142. Marouli E, Graff M, Medina-Gomez C, Lo KS, Wood AR, Kjaer TR, et al. Rare and low-frequency coding variants alter human adult height. Nature. 2017;542(7640):186-90.
143. Okbay A, Beauchamp JP, Fontana MA, Lee JJ, Pers TH, Rietveld CA, et al. Genome-wide association study identifies 74 loci associated with educational attainment. Nature. 2016;533(7604):539-42.
144. Staley JR, Blackshaw J, Kamat MA, Ellis S, Surendran P, Sun BB, et al. PhenoScanner: A database of human genotype-phenotype associations. Bioinformatics (Oxford, England). 2016;32(20):3207-9.
145. Kamat MA, Blackshaw JA, Young R, Surendran P, Burgess S, Danesh J, et al. PhenoScanner V2: An expanded tool for searching human genotype-phenotype associations. Bioinformatics (Oxford, England). 2019;35(22):4851-3.
146. Buniello A, MacArthur JAL, Cerezo M, Harris LW, Hayhurst J, Malangone C, et al. The NHGRI-EBI GWAS Catalog of published genome-wide association studies, targeted arrays and summary statistics 2019. Nucleic acids research. 2019;47(D1):D1005-D12.
147. GTEx Consortium. Human genomics. The Genotype-Tissue Expression (GTEx) pilot analysis: Multitissue gene regulation in humans. Science (New York, NY). 2015;348(6235):648-60.
148. Brion M-JA, Shakhbazov K, Visscher PM. Calculating statistical power in Mendelian randomization studies. International journal of epidemiology. 2012;42(5):1497-501.