Mendelian randomisation study of smoking exposure in relation to breast cancer risk

Hanla A. Park et al.

BACKGROUND: Despite a modest association between tobacco smoking and breast cancer risk reported by recent epidemiological studies, it is still equivocal whether smoking is causally related to breast cancer risk.

METHODS: We applied Mendelian randomisation (MR) to evaluate a potential causal effect of cigarette smoking on breast cancer risk. Both individual-level data as well as summary statistics for 164 single-nucleotide polymorphisms (SNPs) reported in genomewide association studies of lifetime smoking index (LSI) or cigarette per day (CPD) were used to obtain MR effect estimates. Data from 108,420 invasive breast cancer cases and 87,681 controls were used for the LSI analysis and for the CPD analysis conducted among ever-smokers from 26,147 cancer cases and 26,072 controls. Sensitivity analyses were conducted to address pleiotropy.

RESULTS: Genetically predicted LSI was associated with increased breast cancer risk (OR 1.18 per SD, 95% CI: 1.07–1.30, \( P = 0.11 \times 10^{-2} \)), but there was no evidence of association for genetically predicted CPD (OR 1.02, 95% CI: 0.78–1.19, \( P = 0.85 \)). The sensitivity analyses yielded similar results and showed no strong evidence of pleiotropic effect.

CONCLUSION: Our MR study provides supportive evidence for a potential causal association with breast cancer risk for lifetime smoking exposure but not cigarettes per day among smokers.
and 35,314 controls (http://ccge.medschl.cam.ac.uk/research/consortia/iCOGS/). Genotype data were imputed based on the 1000 Genomes project Phase 3 as the reference panel using the programme IMPUTE2.19 SNPs with high imputation quality (imputation \( r^2 \geq 0.5 \)) were included. Overlapping participants between datasets were excluded from the iCOGS dataset as the OncoArray provides a better genomic coverage than the iCOGS array. Demographic and epidemiologic data were harmonised across BCAC sites based on a standardised protocol and derived with respect to a reference date, which was date at diagnosis for cases and date at interview for controls. For controls and cases from the nested case-control studies, data from the baseline interview were considered, or if available, follow-up information. Chosen characteristics of the two datasets, including self-reported smoking behaviours (e.g., smoking status, smoking heaviness and smoking duration, including age at smoking initiation and lifetime smoking exposure) are shown in Supplementary Table 1. Ethics approval was obtained from the relevant institutional review boards for all BCAC studies, and all participants provided written informed consent.

Selection of SNPs associated with smoking exposure

Exposure variables were selected based on the availability of associated genetic variants to reflect smoking exposure. We included two quantitative smoking behaviour-related traits, cigarettes per day (CPD; average number of cigarettes smoked per day by ever-smokers) and lifetime smoking index (LSI; composite score that captures lifetime smoking exposure by taking into account smoking status as well as smoking duration, heaviness and cessation in ever-smokers).20 The recent GWAS of cigarettes per day from GSCAN consortium identified 55 conditionally independent genome-wide significant SNPs, explaining 1.09% of the variance in a sample of 337,334 ever-smokers of European ancestry.21 For LSI, 126 significantly associated SNPs, 0.36% explained variance in LSI, were identified based on a sample of 426,690 individuals (never-smokers and ever-smokers) of European ancestry from the UK biobank.20 The genetic scores of the two smoking behaviour-related traits were reported to be associated with significantly higher risks of lung cancer.20,22

For our analysis, SNPs were selected if they were reported to be associated at genome-wide significance level (\( P \leq 5 \times 10^{-8} \)) and had a minor allele frequency (MAF) above or equal 1%. For each behaviour phenotype, we filtered the list of behaviour-associated SNPs so that the remaining SNPs were not in linkage disequilibrium (LD) (\( r^2 > 0.1 \)). For CPD, the SNP with lowest \( P \) value (rs10519203) of the correlated SNPs (rs12438181, rs28438420, rs72740955, rs146009940, rs28618284, rs804086, \( r^2 > 0.01 \)) was retained. One SNP (rs4886550) was not available in our data and without any proxy SNPs (LD, \( r^2 > 0.8 \)). Given that alcohol consumption is a recognised confounder of the association between smoking and breast cancer risk,23–25 we excluded nine SNPs (three of CPD, six of LSI, respectively) correlated with any alcohol consumption-associated SNPs (\( r^2 > 0.1 \)) from recent large-scale GWAS (drinks per week, \( P \leq 5 \times 10^{-8} \)).22 After exclusions, we included a total of 164 variants associated with one of the smoking traits; 44 and 120 variants, respectively, for CPD and LSI (Supplementary Table 2). No variants overlapped between the two smoking traits.

Statistical methods

Statistical power. Power calculations were conducted to estimate the magnitude of effects detectable with our study size assuming 5% a level and an \( R^2 \) of 0.0109 for CPD and \( R^2 \) of 0.0036 for LSI, which correspond to the variance in each smoking behaviour explained by the SNPs used for this analysis. Power calculations were performed using an online tool available at http://cnsgenomics.com/shiny/mRnd/.26 Detailed power calculations for all outcomes (invasive breast cancer, ER-positive and ER-negative) to detect different odd ratios are shown in Supplementary Table 3.

wPGS-based analyses. For our primary analysis, we generated weighted polygenic scores (wPGSs) using individual-level data of BCAC participants as follow: \( wPGS = \sum_{i=1}^{n} \beta_{gi} a_i \), which is the sum of the effect allele dosage (\( a_i \)) ranging from 0 to 2 for each SNP weighted by the \( \beta \)-coefficient \( \beta_{gi} \) for the effect of the genetic variant (\( g \)) related to the smoking quality or/and duration (CPD, and LSI) (x) (Supplementary Table 2).26,27 The wPGSs of LSI was weakly correlated with wPGSs of CPD (0.17 in iCOGS and 0.03 in OncoArray) in ever-smokers. Analysis of LSI was conducted using all participants (including ever- and never-smokers), whereas that of CPD was performed solely in ever-smokers, and based on 26,147 cases (7342 for iCOGS, 18,805 for OncoArray) and 26,072 controls (8489 for iCOGS, 17,583 for OncoArray). These inclusion criteria correspond to those used in the GWAS studies that identified the SNPs for the smoking traits.20,21 Association analysis of the wPGSs with breast cancer risk using logistic regression was performed using fixed effect meta-analyses combining iCOGS and OncoArray results based on heterogeneity evaluated by Cochran’s Q statistics. The basic model (Model 1) was adjusted for age (continuous), principal components (PCs) of genetic ancestry (first ten PCs for iCOGS and OncoArray, separately) and study site, as previously described.22 In order to assess if the genetic instrument is independent of established risk factors for breast cancer, associations of wPGSs with selected breast cancer risk factors were assessed by linear regression for continuous variables and logistic regression for categorical variables (Supplementary Table 4). We adjusted for the risk factors that were associated with wPGSs of at least one of the smoking traits in Model 2. We additionally adjusted for alcohol assumption, a well-known confounder of the association between breast cancer risk and smoking, separately in Model 3 because of the large amount of missing data. Participants with missing covariates were excluded from all of the analyses. Stratified analysis was performed to assess potential differences in associations with breast cancer risk by menopausal status (pre-, postmenopausal women) adjusting for age, study site and top ten PCs. Heterogeneity was tested employing the likelihood ratio test (LRT) for evaluating the multiplicative interaction terms in nested models. Polytomous regression was used to estimate the association according to oestrogen receptor (ER) status.23–25 Scaling was applied to convert the wPGSs for CPD into meaningful units through dividing them by linear regression coefficients of self-reported CPD (0.35 per pack of cigarettes per day). The regression coefficient of CPD was derived from a meta-analysis of iCOGS and OncoArray data on smoking behaviours among 26,072 among ever-smoker controls (8490 and 17,703 controls of iCOGS and OncoArray, respectively).

Two-sample MR analyses. Five different two-sample MR methods using summary association data were applied: inverse-variance weighted (IVW),28 MR Egger,29 weighted median,30 weighted mode,31 and robust adjusted profile score (RAPS).32 Each of these methods makes slightly different assumptions about the nature of pleiotropy and therefore a roughly consistent odd ratios are provided. The IVW method was implemented since the instruments consisted of multiple SNPs.33 Multivariable MR methods34 were conducted also using summary association data from GWASs of alcohol consumption (drinks per week),35 body mass index (BMI) among females,35 and education attainment.36 To produce valid results, the IVW method requires that all instruments are associated with the exposure of interest (relevance assumption), but neither directly with the outcome of interest (only via the exposure; exclusion restriction) nor any confounders of the relationship between the exposure and the outcomes.
outcome (independence assumption). The intercept from MR-Egger regression is a statistical test for horizontal pleiotropy, whereas the slope can be interpreted as the smoking behaviour effect on breast cancer adjusted for horizontal pleiotropy. This method assumes however that the pleiotropic effects are independent of the instrument strength (InSIDE assumption). The weighted median estimator provides a valid causal estimate when at least half of the instruments are valid. MR-RAPS test extends the basic IVW random-effects approach by making the weight each variant receives in the analysis a function of the causal effect and the precision of the SNP-exposure association. The MR pleiotropy residual sum and outlier test (MR-PRESSO) was also implemented to identify outlying genetic variants and analyses were re-run after excluding these variants.

All two-sample MR analyses using summary association data were performed with respect to three cancer susceptibility phenotypes: overall breast cancer (108,067 cases/88,386 controls) as well as oestrogen receptor (ER)-positive (70,435 cases) and ER-negative tumours (17,365 cases). Due to the nature of summary-level data, the analyses of both LSI and CPD were conducted using all samples regardless of smoking status.

R version 3.4.3 was used to conduct analyses. R package "Mendelian randomisation", "mr_raps" and "MR-PRESSO" were used for two-sample MR analysis. All tests were considered at the 0.05 level of significance.

RESULTS

There was an association of wPGS for LSI with increased invasive breast cancer risk (OR per SD 1.18, 95% CI: 1.07–1.30, P = 0.11 × 10⁻²) whereas little evidence was found for an association between wPGSs for CPD and invasive breast cancer (OR 1.02 per pack of cigarettes per day, 95% CI: 0.78–1.19, P = 0.85) after adjustment for age and study (Model 1) (Table 1). Several breast cancer risk factors were associated with wPGSs of one of the smoking traits (CPD or/and LSI), including ever breastfeeding, menopausal status, age at menopause, BMI, age at first live birth, parity and education level (Supplementary Table 4). Adjustment for all of the identified risk factors did not change the association substantially (Model 2) (wPGSs for LSI, OR per SD 1.24, 95% CI: 1.06–1.45, P = 0.60 × 10⁻²) (Table 1). The point estimate of association between wPGS for LSI and invasive breast cancer risk remained unchanged after additional adjustment for alcohol consumption although imprecisely estimated (i.e. wide confidence intervals) (Model 3) (OR per SD 1.13, 95% CI: 0.86–1.49, P = 0.39).

The association between wPGSs for CPD and invasive breast cancer did not change after adjustments.

There was no evidence for effect heterogeneity of the associations of LSI and CPD with breast cancer risk according to ER status or menopausal status (Supplementary Fig. 1). Using IVW random-effects analysis, positive associations of genetically predicted LSI were found for overall breast cancer risk (OR per SD 1.14, 95% CI: 1.02–1.28, P = 0.02) and breast cancers according to ER status (OR per SD 1.14, 95% CI: 1.00–1.30, P = 0.04, for ER-positive and OR per SD 1.14, 95% CI: 0.95–1.37, P = 0.17, for ER-negative tumours) (Table 2, Supplementary Table 5 and Supplementary Figs. 2–4). There was no indication of horizontal pleiotropy based on the MR-Egger intercept test for any outcomes. The point estimates of associations were consistent across the different methods although the results based on MR-Egger regression, and weighted mode method were imprecisely estimated (i.e. wide confidence intervals). They remained substantially unchanged after multivariable adjustment for alcohol consumption, BMI and education. The MR-PRESSO analysis revealed three outliers for LSI. Removal of an outlier (rs2867112) with respect to risk for overall breast cancer and ER-negative cancer did not change the associations. No outlier was observed with ER-positive tumour (Table 2, Supplementary Table 5 and Supplementary Figs. 2–4).

For genetically predicted CPD, we found little evidence for an association with overall breast cancer and ER subtypes using IVW random-effects method (Table 2, Supplementary Table 5 and Supplementary Figs. 2–4). There was no indication of horizontal pleiotropy based on the MR-Egger intercept test. Results across two-sample MR analyses were consistent in overall breast cancer and ER subtypes without any indication of horizontal pleiotropy from the MR-Egger intercept. The MR-PRESSO analysis detected three outlying SNPs (rs11940255, rs1737894 and rs73229090) although the outcome of IVW analysis after outlier removal remained unchanged in overall breast cancer and ER-positive tumour risk. No outlier was observed with ER-negative disease (Table 2, Supplementary Table 5 and Supplementary Figs. 2–4).

DISCUSSION

This MR study supports an association between genetically predicted lifetime smoking exposure and increasing invasive breast cancer risk but no clear association with cigarettes per day among smokers. The estimates based on the wPGS for the two smoking traits and several two-sample MR methods were consistent.

Table 1. Associations of genetic risk scores for cigarette smoke exposure-related traits with breast cancer risk: results from Mendelian randomisation analysis.

|                | CPD | LSI          |
|----------------|-----|--------------|
|                | No. of cases/controls OR 95% CI P | No. of cases/controls OR 95% CI P |
| Model 1a       | 26,147/26,072 1.02 0.78–1.19 0.85 | 108,420/87,681 1.19 1.07–1.30 1.1 × 10⁻³ |
| Model 2b       | 7360/7168 0.98 0.74–1.33 0.95 | 17,936/16,654 1.24 1.06–1.45 6.0 × 10⁻³ |
| Model 3c       | 2892/2754 0.66 0.36–1.19 0.16 | 7716/7028 1.13 0.86–1.49 0.39 |

No. number, CPD cigarettes per day, LSI lifetime smoking index, OR odds ratio per year, CI confidence interval, P P value, BMI body mass index.

*Adjusted for age, sex and top ten PCs. P value for heterogeneity between iCOGS and OncoArray data of cigarettes per day and lifetime smoking index are 0.60, 0.09, respectively; *in addition to adjustment of Model 1, additionally adjusted for ever breastfeeding, postmenopausal status, age at menopause, BMI, age at first live birth, parity and education level. P value for heterogeneity between iCOGS and OncoArray data of cigarettes per day and lifetime smoking index are 0.96, 0.37, respectively; *in addition to adjustment of Model 2, additionally adjusted for alcohol consumption (glasses per day) P value for heterogeneity between iCOGS and OncoArray data of cigarettes per day and lifetime smoking index are 0.80, 0.66 respectively; *OR per pack of cigarettes per day; *OR per standard deviation.
The LSI has not been assessed in studies of breast cancer risk but the modest association found in this analysis is in line with the modest associations of current and former smoking with invasive breast cancer risk reported in the recent large pooled analysis.\textsuperscript{13} We did not find support for an association between cigarettes per day in ever-smokers and invasive breast cancer risk, whereas modest associations were reported for cigarettes per day in current smokers compared with never-smokers in the pooled analysis. With the restriction to ever-smokers, the MR analysis of CPD had low statistical power for the very modest dose–response association estimated by our data (Table 1) as well as that reported in the pooled analysis.\textsuperscript{13}

In addition to smoking exposure traits, we addressed two dichotomous smoking status traits, namely smoking initiation and smoking cessation, which account for the lifetime smoking exposure. Despite a pooled analysis of epidemiological studies reporting an increased risk of breast cancer associated with current smokers compared to non-smokers (OR of 1.02),\textsuperscript{13} a recent MR study found inconclusive evidence of association between genetically predicted smoking initiation and breast cancer risk using summary-level data (OR: 1.05, 95% CI: 0.99–1.12, \(P = 0.12\)) in BCAC; OR: 0.97, 95% CI: 0.90–1.06, \(P = 0.51\) in UK Biobank).\textsuperscript{39} We conducted association analyses of wPGS of smoking initiation and smoking cessation with breast cancer risk (see Supplementary Note) and found no statistically significant association for either smoking initiation (OR 1.05, 95% CI 0.99–1.11, \(P = 0.08\)) or smoking cessation (OR 1.06, 95% CI: 0.88–1.27, \(P = 0.52\)) (Supplementary Table 6). The results remained unchanged after adjusting for breast cancer risk factors (Supplementary Table 6). We cannot rule out that the MR analysis of the smoking status had low statistical power to detect the very modest association estimated by our data (Table 1) along with that reported in the pooled analysis.\textsuperscript{13} It is also possible that the MR analysis of smoking status especially smoking initiation alone does not capture the association between cigarette smoking and breast cancer risk comparing to the LSI which accounts for other smoking traits.

There are several not entirely resolved issues concerning the association between smoking and breast cancer risks, such as potential effect modification by timing of smoking exposure, menopausal status and oestrogen receptor (ER) status, potential confounding by alcohol consumption. We conducted an association analysis between wPGS of age at smoking initiation and invasive breast cancer (see Supplementary Note). The result showed an inverse association but statistically nonsignificant with low precision (OR 0.88, 95% CI: 0.25–3.07, \(P = 0.84\)) (Supplementary Table 6). This is in line with the result of a pooled analysis of epidemiological studies showing that women who started smoking later than 24 years old were at lower breast cancer risk than those who started smoking earlier when compared to non-smokers.\textsuperscript{13} Smoking initiation in relation to first birth has been considered an essential factor in the association with breast cancer risk since the undifferentiated breast epithelium is particularly susceptible to carcinogens before the first birth.\textsuperscript{40} Indeed, this appears to be supported by findings of a stronger association with smoking with breast cancer risk if initiated before first birth and a stronger influence of smoking on breast cancer among women who started smoking more than 10 years before the first full-term pregnancy on breast cancer.\textsuperscript{13,41–46} Since the relevant information was only available for a subset of study participants, we did not have sufficient power to address potential differential associations according to timing of smoking exposure in relation to first birth.

Stronger associations between smoking and breast cancer among premenopausal women have been hypothesised since the morphology of the breast and the endogenous hormone levels change substantially during the menopausal transition, and menopausal status alter other breast cancer risk factors.\textsuperscript{10} The MR results confirmed the lack of effect modification by menopausal status also reported by previous epidemiological studies.\textsuperscript{13,47,48} Despite early evidence against a differential association by ER status,\textsuperscript{24,49,50} recent epidemiologic studies reported a stronger association for risk of ER-positive breast cancer.\textsuperscript{10,13,41} We did not find clear evidence for heterogeneity by ER status although power to detect effect heterogeneity was limited particularly due to the small sample size for ER-negative disease.

---

**Table 2.** Association of cigarette smoke with overall breast cancer risk: results from the two-sample Mendelian randomisation using summary statistics.

| CPD | LSI |
|-----|-----|
| IVW summary statistics\textsuperscript{a} | OR\textsuperscript{a} (95% CI) | OR\textsuperscript{a} (95% CI) |
| 1.02 | 0.89–1.17 | 0.74 | 1.14 | 1.02–1.28 | 2.5 × 10\textsuperscript{-02} |
| Multivariable IVW summary statistics\textsuperscript{a,b} | 1.05 | 0.92–1.20 | 0.49 | 1.38 | 1.08–1.76 | 8.9 × 10\textsuperscript{-03} |
| Estimate using IVW after outlier corrected\textsuperscript{a,c} | 1.01 | 0.88–1.16 | 0.87 | 1.17 | 1.05–1.31 | 6.2 × 10\textsuperscript{-03} |
| MR-Egger regression\textsuperscript{a,d} | 0.98 | 0.77–1.24 | 0.84 | 1.39 | 0.88–2.20 | 0.15 |
| MR multivariable Egger regression\textsuperscript{k,k,w} | 1.03 | 0.82–1.29 | 0.82 | 1.36 | 0.88–2.11 | 0.17 |
| MR weighted median estimator\textsuperscript{a,f} | 1.06 | 0.93–1.20 | 0.40 | 1.17 | 1.02–1.33 | 2.1 × 10\textsuperscript{-02} |
| MR robust adjusted profile score estimator\textsuperscript{a,g} | 1.02 | 0.94–1.12 | 0.57 | 1.44 | 1.32–1.58 | 4.7 × 10\textsuperscript{-15} |
| MR weighted mode\textsuperscript{a,h} | 1.01 | 0.89–1.15 | 0.84 | 1.23 | 0.86–1.75 | 0.26 |

\textsuperscript{a}OR odd ratio, MR Mendelian randomisation, IVW inverse-variance weighted, CI confidential interval, CPD cigarettes per day, LSI lifetime smoking index, RAPS robust adjusted profile score.

All two-sample MR analyses using summary-level data were performed in all samples regardless of smoking status (108,067 overall breast cancer cases/88,386 controls); \textsuperscript{a}estimate derived using summary statistics (28); \textsuperscript{b}multivariable analysis after adjusting for genetically predicted alcohol consumption (drinks per week), body mass index and education attainment by using summary-level data from GWAS outcome (alcohol assumption (21), body mass index (BMI) among female (35) and education attainment (36); \textsuperscript{c}the MR pleiotropy residual sum and outlier test (MR-PRESSO) was implemented to identify outlying genetic variants (rs11940255, rs1737894 and rs73229090 for CPD; rs2867112 for LSI) and analyses were re-run after excluding these variants (38); \textsuperscript{d}the MR-Egger intercept yielded no indication of strong pleiotropic effects (CPD: \(\beta = 0.24\pm 0.04, P = 0.99\); LSI: \(\beta = -3.14\pm 0.37\)); \textsuperscript{e}the multivariable MR-Egger intercept yielded no indication of strong pleiotropic effects (CPD: \(\beta = 6.88\pm 0.04, P = 0.83\); LSI: \(\beta = 3.24\pm 0.93\)); \textsuperscript{f}estimates derived using weighted median estimator approach, where 50% of the variants included in each genetic instrument are assumed to be invalid (30); \textsuperscript{g}estimates derived using Mendelian randomisation robust adjusted profile score (MR-RAPS) method (32); \textsuperscript{h}estimates derived using weighted mode estimator approach, where the largest group of instruments with consistent MR estimate are assumed to be valid (31); \textsuperscript{i}OR per standard deviation.
Previous epidemiologic studies have addressed the confounding effect by alcohol consumption and found an association of smoking with breast cancer risk after stratifying on alcohol consumption.13,41–43 We addressed this issue by excluding SNPs associated with alcohol intake from the wPGS. However, the results did not change significantly in sensitivity analyses with wPGS, including the overlapping alcohol consumption-associated SNPs (Supplementary Table 7). Also, our multivariable analyses adjusting for alcohol consumption yielded association estimates that remained unchanged. The lower precision can be attributed to the reduced dataset that required information on alcohol intake with the ensuing diminished power.

Despite epidemiological evidence indicating alcohol consumption as an established risk factor for breast cancer risk,49 a recent MR study reported no significant association of breast cancer risk with genetically predicted alcohol intake using summary association data.39 We conducted an analysis between wPGS of alcohol consumption and invasive breast cancer using individual-level data (see Supplementary Note). We found no clear evidence of association between genetically predicted alcohol consumption and invasive breast cancer (OR: 0.98, 95% CI: 0.86–1.11, P = 0.74). The association remained unchanged after adjusting for LSI (OR: 0.88, 95% CI: 0.72–1.09, P = 0.25).

Strengths and limitations
Despite the relatively large sample size used for our analyses, weak statistical power to detect modest associations should be considered when interpreting the results. The power calculation shows that our study had 68% power to detect an OR of 1.20 per SD change in LSI but only around 9% power for an OR of 1.05 in combined dataset (iCOGS and OncoArray) (Supplementary Table 3). Since the reported order of magnitude for associations with CPD is around 1.05 per ten cigarettes,13 we cannot exclude that the lack of association observed in the MR analysis particularly regarding CPD may be due to limited power. The unit for LSI is not scaled due to the nature of the phenotype. The reported order of magnitude for associations with LSI can therefore not be directly compared with the relative risk estimates for smoking from observational studies.

MR estimates have a causal interpretation only if the assumptions of the instrumental variable approach hold. Even though we performed extensive sensitivity analyses to detect potential violations, it is difficult to prove the validity of the assumptions. The LSI captures multiple aspects of smoking behaviours, which could have introduced more potential for horizontal pleiotropy. More the diffusion the definition of smoking, more the lifestyle factors might be correlated, making it especially important to test for horizontal pleiotropy. No evidence of pleiotropic effects was found by conducting various sensitivity analyses; however, residual pleiotropy is difficult to exclude and should be considered.

The genetic instrument for LSI allows for the use of the large entire sample to conduct MR analysis without stratifying on smoking status. The analyses for CPD were restricted to smokers by reason that the CPD-associated SNPs were identified among ever-smokers in the GWAS study,13 which reduce statistical power to detect an association. Moreover, we should note that restricting to ever-smokers may induce a sampling bias and invalidate the MR assumptions. By restricting to smokers, smoking initiation can open up the path from exposure (wPGS for CPD) to outcome (breast cancer risk). It can make the association between smoking and breast cancer risk appear weaker by removing a part of the association that is attributable to smoking initiation.

Another limitation is that our analysis was restricted to participants of European ancestry; therefore, our results may not apply to populations of other ethnicities. However, it reduces the potential bias caused by population stratification.

Conclusion
In conclusion, this Mendelian randomisation analysis using both individual-level data and summary statistics supports a causal association between lifetime smoking exposure and breast cancer risk. Larger studies for MR analysis are warranted to address additional aspects of smoking behaviour.
We thank all the individuals who took part in these studies and all the researchers, clinicians, technicians and administrative staff who have enabled this work to be carried out. The COGS study would not have been possible without the contributions of the following: Andrew Berchuck (OCAC), Rosalind A. Eeles, Ali Amin Al Olama, Zsofia Kote-Jarai, Sara Benlloch (PRACTICAL), Antonion Antoniou, and Lesley McCugфф (CIBMB). Andrew Lee, and Ed Dicks, Craig Luccarini and the staff of the Centre for Genetic Epidemiology Laboratory, Anna Gonzalez-Neira and the staff of the CNIO genotyping unit, and Daniel C. Tessier, Francois Bacoet, Daniel Vincent, Sylvie LeBoissière and Frederic Robidoux and the staff of the McGill University and Génome Québec Innovation Centre, Sune F. Nielsen and the staff of the Copenhagen DNA laboratory, and Julie M. Cunningham, Sharon A. Windebank, Christopher A. Hikker, Jeffrey Meyer and the staff of Mayo Clinic Genotyping Core Facility. ABCFS thank Maggie Angelakos, Judi Maskell, Gillian Ditte. ABCS thanks the Blood bank Sanquin, THE NETHERLANDS. ABCCTB Investigators: Christine Clarke, Rosemary Ballestine, Robert Baxter, Stephen Brayne, Jane Carpenter, Jane Dahlstrom, John Forbes, Soon Lee, Debbie Mars, Adrienne Morey, Nirmala Pathmanathan, Rodney Scott, Allan Spigelman, Nicholas Wilcken, Desmond Yip. Samples are made available to researchers on a non-exclusive basis. BCBS thanks Eileen Williams, Elaine Ryder-Mills, Kari Sargus. BCEES thanks Allyson Thompson, Christobel Saunders, Terey Slaven, BreastScreen Western Australia. Elenor, and Fred Schi Lloyd. The COGS study would not have been possible without the contributions of Dr. K. Landsman, Dr. N. Gronich, Dr. A. Flugman, Dr. W. Saliba, Dr. E. Liani, Dr. I. Cohen, Dr. S. Kalet, Dr. V. Friedman, Dr. O. Barette of the NICCC in Haffa, and all the contributing family medicine, surgery, pathology and oncology teams in all medical institutes in Northern Israel. BIGSNUM thanks Nail McNierney, Gabrielle Colleen, Andrew Rowan, Angela Jones. The BRELOGO study would not have been possible without the contributions of the following: Manuela Gago-Dominguez, Jose Esteban Castano, Angel Carracedo, Victor Muñoz Garzón, Alejandro Novo Dominguez, Maria Elena Martinez, Sara Miranda Ponte, Carmen Redondo Maley, Maite Peña Fernández, Manuel Enguíx Castelo, María Torres, Manuel Calaza (BRELOGO), José Antúnez, Máximo Fraga and the staff of the Department of Pathology and Biobank of the University Hospital Complex of Santiago-CHUS, Instituto de Investigación Sanitaria de Santiago, IDIS, Xerencia de Xestión Integrada de Santiago-SERGAS; Joaquín González-Carreró and the staff of the Department of Pathology and Biobank of University Hospital Complex Virgen del Rocio, Instituto de Investigación Biomédica Galicia Sur, SERGAS, Vigo, Spain. BSUCH thanks Peter Bugert, Medical Faculty Mannheim. CBCS thanks study participants, co-investigators, collaborators and staff of the Canadian Breast Cancer Foundation. Our project coordinators Agnès-Léa Masse, Sophie St-Léonard, Stéphanie Apostolaki, Anna Margioliakis, Georgios Nintos, Maria Perraki, Georgia Vigo, Spain. BCBS thanks Peter Bugert, Medical Faculty Mannheim. CBCS thanks study participants, co-investigators, collaborators and staff of the Canadian Breast Cancer Foundation. Our project coordinators Agnès-Léa Masse, Sophie St-Léonard, Stéphanie Apostolaki, Anna Margioliakis, Georgios Nintos, Maria Perraki, Georgia Vigo, Spain. BCBS thanks Peter Bugert, Medical Faculty Mannheim. CBCS thanks study participants, co-investigators, collaborators and staff of the Canadian Breast Cancer Foundation. Our project coordinators Agnès-Léa Masse, Sophie St-Léonard, Stéphanie Apostolaki, Anna Margioliakis, Georgios Nintos, Maria Perraki, Georgia Vigo, Spain.
Breast Cancer Now and the Institute of Cancer Research for support and funding of the Breakthrough Generations Study, and the study participants, staff and the doctors, nurses and other health care providers and health information sources who have contributed to the study. We acknowledge NHS funding to the Royal Marsden/ICR NIHR Biomedical Research Centre. We thank the UKOPS participants who donated their samples and data and all the researchers, especially Andy Ryan, Simon Gayther and Ian Jacobs.

AUTHOR CONTRIBUTIONS
J.C. conceived the project. H.A.P and J.C-C. conducted analyses, interpreted the results and wrote and edited the original and revised paper. S.N. contributed to the data analyses and the interpretation of the results. S.B.e and P.M.R. contributed to the data preparation. D.F.E., W.Z. and P.K.R. contributed to the results interpretation and the manuscript drafting and revision. D.F.E. coordinated the BCAC and led the iCOGS and OncoArray genotyping. P.H. led the CGGS collaboration, J.S. led the OncoArray collaboration. M.K.B., Q.W. and J.Den. coordinated the BCAC database. K.M.I., T.U.A., I.L.A., A.H.-C., N.N.A., V.A., K.J.A., A.Au., A.B., L.E.B.F., H.Bec., M.W.B., S.B.e, J.Bern., M.Ber., N.V.B., S.E.B., H.Bra., H.Bre., S.Y.B., B.Bur., D.Cam., F.C., J.E.C., J.S.C., G.C.-T., C.L.C., N.C.M., D.C., F.J.C., A.Cox, S.C.C., K.Cz., M.B.D., P.D., T.D., I.D.-S., S.D., M.Dw., D.M., M.E., A.H., Cn., M.E., D.G.E., P.A.F., H.F., L.F., M.-G.C., J.A.G.-S., M.M.G., G.G., M.S.G., D.E.G., A.G.-N., M.G., P.G., E.H., C.A.H., N.H., P.H., U.H., S.H., E.F.H., N.S.H., W.H., B.A.M.H.G., J.L.H., D.J., A.I., A.J., A.I., A.J., R., P.M.K., R.K., E.K., C.M.Ki., L.B.K., S.Ki., V.N.K., Y.W.K., A.K.W., J.V.L., D.La., L.M., W.-L., J.Ju., A.Man., M.Man., M.Mar., M.E.M., D.M., A.Me., U.Me., R.L.M., T.A.M., H.Ne., W.G.N., B.G.N., K.O.F., A.F.O., H.O., T.W.-P.S., P.F., P.E., J.P., D.P.K., N.P., P.R.a., G.R., H.S.R., A.R.o., E.S., J.I.S., M.K.S., R.K.S., M.J.S., L.Sch., C., M.Sh., X.O.S., J.Si., A.Sm., M.C.S.I., J.J.S., V.S., A.I.S., R.M.T., W.I.T., I.A.T., M.B.T., I.A.T., M.T., C.M.T., C.V., E.M.W., J.V.i., S.W., C.W.e., R.W., A.W., A.Z., P.D.P., W.Z., P.K. and J.C.C. led individual studies and contributed to the design of the study and data collection.

ADDITIONAL INFORMATION
Ethics approval and consent to participate Collection of blood samples, urine samples and questionnaire information was undertaken with written informed consent and relevant ethical review board approval in accordance with the tenets of the Declaration of Helsinki (Supplementary Table 8).

Consent to publish Not applicable.

Data availability Availability of lifetime smoking index GWAS described by Wootton et al.20 Availability of self-reported cigarettes per day GWAS described by Liu et al.21 Individual genotyping data from BCAC will not be made publicly available due to restraints imposed by the ethics committees of individual studies; requests for data can be made to the Data Access Coordination Committee of BCAC (http://bcac.ccge.medschl.cam.ac.uk). Summary results for all variants genotyped by BCAC are available at http://bcac.ccge.medschl.cam.ac.uk/.

Competing interests U.Me. has reported a conflict of interest outside of the submitted work (owing stocks in Abcodia awarded to her by University College London). P.H., R.M.T. and A.Cox. are subject editors for the British Journal of Cancer. P.D.P., E.J.S. and K.Cz. are editorial board members for the British Journal of Cancer. The remaining authors declare no competing interests.

Funding information BCAC is funded by Cancer Research UK [C1287/A16563, C1287/A10118], the European Union's Horizon 2020 Research and Innovation Programme (grant numbers 634935 and 633784 for BRIDGES and B-CAST respectively), and by the European Community's Seventh Framework Programme under grant agreement number 223175 (grant number HEALTH-F2-2009-223175) (COGS). The EU Horizon 2020 Research and Innovation Programme funding source supported the Government of Canada through Genome Canada and the Canadian Institutes of Health Research (grant GPH-129344) and, the Ministère de l'Education Nationale, Institut de l'Enfance et de la Santé et de la Recherche Médicale (INSERM) (France); German Cancer Research Center (DKFZ). CBCS is funded by the dutch Dutch Cancer Society [grants NKI 2007-3839; 2009 4363]. The Australian Breast Cancer Tissue Bank (ABCt) was supported by the National Health and Medical Research Council of Australia, The Institute of Cancer Research, and the Ovarian Cancer Research Fund. The DRIVE Consortium was funded by U19 CA148065. The Australian Breast Cancer Family Study (ABCFS) was supported by grant U19 CA164920 from the National Cancer Institute (USA). The context of this manuscript does not necessarily reflect the views or policies of the National Cancer Institute or any of the collaborating centres in the Breast Cancer Family Registry (BCFR), nor does mention of trade names, commercial products, or organisations imply endorsement by the USA Government or the BCFR. The ABCFS was also supported by the National Health and Medical Research Council of Australia, the New South Wales Cancer Council, the Victorian Health Promotion Foundation (Australia) and the Victorian Breast Cancer Research Consortium. J.L.H. is a National Health and Medical Research Council (NHMRC) Senior Principal Research Fellow. M.C.S. is a NHMRC Senior Research Fellow. The ABCS study was supported by the Dutch Cancer Society [grants NKI 2007-3839; 2009 4363]. The Australian Breast Cancer Foundation (ABCt) was supported by the National Health and Medical Research Council of Australia, The Institute of Cancer Research, and the Ovarian Cancer Research Fund. The AHS study is supported by the intramural research program of the National Institutes of Health, the National Cancer Institute (grant number Z01-C00119), and the National Institute of Environmental Health Sciences (grant number Z01-ES049030). The work of the BBCC was partly funded by ELAN-Fond of the University Hospital of Erlangen. The BBCS is funded by Cancer Research UK and Breast Cancer Now and acknowledges NHS funding to the NIHR Biomedical Research Centre, and the National Cancer Research Network (NCRN). The BCCEFS was funded by the National Health and Medical Research Council, Australia. For the BCFR-NY, BCFR-PA, BCFR-UT, this work was supported by grant U116 CA164920 from the National Cancer Institute. The context of this manuscript does not necessarily reflect the views or policies of the National Cancer Institute or any of the collaborating centres in the Breast Cancer Family Registry (BCFR), nor does mention of trade names, commercial products, or organisations imply endorsement by the US Government or the BCFR. For BIGGS, ES is supported by NIHR Comprehensive Biomedical Research Centre, Guy's & St Thomas' NHS Foundation Trust in partnership with King's College London, United Kingdom. It is supported by the Oxford Biomedical Research Centre. The Breast Oncology Galician Network (BREGON) is funded by Acción Estratégica de Salud del Instituto de Salud Carlos III FIS PI12/02125/Cofinanciado FEDER; Acción Estratégica de Salud del Instituto de Salud Carlos III FIS Intraisl (PI13/01336); Programa Grupos Emergentes, Cancer Genetics Unit, Instituto de Investigacion Biomedica Galicia Sur. Xerencia de Xestion Integrada de Vigo-SEGRAS, Instituto de Salud Carlos III, Spain; Grant 102012/2, CEI, University Foundation of Investigation Applied, PHEME I + D e I + D-Suma del Plan Gallego de Investigación, Desarrollo e Innovación Tecnológica de la Conselleria de Industria de la Xunta de Galicia, Spain; Grant EC11-192. Fomento de la Investigación Clínica Independiente, Ministerio de Sanidad, Servicios Sociales e Igualdad, Spain; and Grant FEDER-Innterconecta. Ministerio de Economía y Competitividad, Xunta de Galicia, Spain. The BSUCH study was supported by the Dietmair-Hopp Foundation, the Helmholtz Society and the German Cancer Research Center (DKFZ). CBS is funded by the Canadian Cancer Society (grant # 313404) and the Canadian Institutes of Health Research. CCGP is supported by funding from the University of Crete. The CECILE study was supported by Fondation de France, Institut National du Cancer (INCa), Ligue Nationale contre le Cancer, Agence Nationale de Sécurité Sanitaire, de l’Alimentation, de l’Environnement et du Travail (ANSES), Agence Nationale de la Recherche (ANR). The CCGP was supported by the Uehlian-Physician John Mayos Research Fund, the Danish Medical Research Council, and Herlev and Gentofte Hospital. The CNIO-BCS was supported by the Instituto de Salud Carlos III, the Red Temática de Investigación Cooperativa en Cáncer and grants from the Asociación Española Contra el Cáncer and the Fondo de Investigación Sanitario (PI11/00923 and PI12/00070). The American Cancer Society funds the creation, maintenance, and updating of the CPS-II cohort. This was initially supported by the California Breast Cancer Act of 1993 and the California Breast Cancer Research Fund (contract 97-10500) and is currently funded through the National Institutes of Health (R01 CA77398, UM1 CA146917 and U01 CA199277). Collection of cancer incidence data was supported by the California Department of Public Health as part of the statewide cancer reporting program mandated by California Health and Safety Code Section 103885. The University of Westminster curates the Dietcomplf database funded by Against Breast Cancer Registered Charity No. 1121258 and the NCRN. The coordination of EPIC is financially supported by the European Commission (DG-SANCO) and the International Agency for Research on Cancer. The national cohorts are supported by: Ligue Contre le Cancer, Institut Gustave Roussy, Mutuelle Générale de l'Education Nationale, Institut National de la Santé et de la Recherche Médicale (INSERM) (France); German Cancer Aid (Deutsche Krebshilfe (DKH), Federal Ministry of Education and Research (BMBF) (Germany); the Hellenic Health Foundation, the Stavros Niarchos Foundation (Greece); Associazione Italiana per la Ricerca sul Cancro-AIRC-Italy and National Research Council (Italy); Dutch Ministry of Public Health, Welfare and Sports (VWS), Netherlands Cancer Registry (RNIK), UK Research Funds, Dutch Prevention Funds, Dutch ZON (Zorg Onderzoek Nederland), World Cancer Research Fund (WCRF), Statistics Netherlands (The Netherlands); Health Research Fund (FIS), PI13/1141
Mendelian randomisation study of smoking exposure in relation to breast... HA. Park et al.
REFERENCES

1. Bray, F., Ferlay, J., Soerjomataram, I., Siegel, R. L., Torre, L. A. & Jemal, A. Global cancer statistics 2018: GLOBOCAN estimates of incidence and mortality worldwide for 36 cancers in 185 countries. CA Cancer J. Clin. 68, 394–424 (2018).

2. Tobacco smoke and involuntary smoking, IARC Monographs on the Evaluation of Carcinogenic Risks to Humans. 83, 1–1438 (2004).

3. Centers for Disease C, Prevention, National Center for Chronic Disease P, Health P, Office on S, Health. Publications and reports of the surgeon general. in How Tobacco Smoke Causes Disease: The Biology and Behavioral Basis for Smoking-Attributable Disease: A Report of the Surgeon General Centers for Disease Control and Prevention (US): Atlanta (GA) (2010).

4. Hecht, S. S. Tobacco smoke carcinogens and breast cancer. Environ. Mol. Mutagen. 39, 119–126 (2002).

5. Petrakis, N. L., Gruenke, L. D., Beelen, T. C., Castagnoli, N. Jr. & Craig, J. C. Nicotine in breast fluid of nonlactating women. Science 199, 303–305 (1978).

6. Petrakis, N. L., Maack, C. A., Lee, R. E. & Lyon, M. Mutagenic activity in nipple aspirates of human breast fluid. Cancer Res. 40, 188–189 (1980).

7. Conway, K., Edmiston, S. N., Cui, L., Drouin, S. S., Pang, J., He, M. et al. Prevalence and spectrum of p53 mutations associated with smoking in breast cancer. Cancer Res. 62, 1987–1995 (2002).

8. Li, D., Zhang, W., Sahin, A. A. & Hittelman, W. N. DNA adducts in normal tissue adjacent to breast cancer: a review. Cancer Detect. Prev. 23, 454–462 (1999).

9. Rundle, A., Tang, D., Hibshoosh, H., Estabrook, A., Schnabel, F., Cao, W. et al. The relationship between genetic damage from polycyclic aromatic hydrocarbons in breast tissue and breast cancer. Carcinogenesis 21, 1281–1289 (2000).

10. Warren, G. W., Alberg, A. J., Kraft, A. S. & Cummings, K. M. The 2014 Surgeon General’s report: “the health consequences of smoking—50 years of progress”: a paradigm shift in cancer care. Cancer 120, 1914–1916 (2014).

11. Band, P. R., Le, N. D., Fang, R. & Deschamps, M. Carcinogenic and endocrine disrupting effects of cigarette smoke and risk of breast cancer. Lancet 360, 1044–1049 (2002).

12. Clemons, M. & Goss, P. Estrogen and the risk of breast cancer. New Engl. J. Med. 344, 276–285 (2001).

13. Gauted, M., Carter, B. D., Brinton, L. A., Falk, R. T., Gram, I. T., Luo, J. et al. Pooled analysis of active cigarette smoking and invasive breast cancer risk in 14 cohort studies. Int. J. Epidemiol. 46, 881–893 (2017).

14. Smith, G. D. & Ebrahim, S. Mendelian randomization: can genetic epidemiology contribute to understanding environmental determinants of disease? Int. J. Epidemiol. 32, 1–22 (2003).

15. Evans, D. M. & Davey Smith, G. Mendelian randomization: new applications in the paradigm of pleiotropy in causal relationships inferred from Mendelian randomization. J. Cancer 7, 25–30 (2016).

16. Murray, C. J. L., Lopez, A. D., Mathers, C. D., Ezzati, M., Lim, S. S., Shibuya, K. et al. Disability-adjusted life years (DALYs) for 89 diseases and injuries 1990–2010: a systematic analysis for the Global Burden of Disease Study 2010. Lancet 387, 2197–2223 (2011).

17. World Health Organization. (2016). Tobacco or health: choosing not to be smoked. http://www.who.int/tobacco/global_report/en/ (2016).

18. World Health Organization. The Global Report on Tobacco Control, 2013. http://www.who.int/tobacco/global_report/2013/en/ (2013).

19. Park, H. A., Park, M. G., Park, Y., Lee, S., Park, J. Y., Jee, S. Y. et al. Smoking, alcohol use, and breast cancer: a Mendelian randomization study in the single-sample and two-sample summary-data Mendelian randomization using robust adjusted profile score. Annals of Statistics, 48, 1742–1769 (2020).

20. Burgess, S., Butterworth, A. & Thompson, S. G. Mendelian randomization analysis with multiple genetic variants using summarized data. Genetic Epidemiol. 37, 658–665 (2013).

21. Sanderson, E., Davey Smith, G., Windmeijer, F. & Bowden, J. An examination of multivariable Mendelian randomization with some invalid instruments using a weighted median estimator. Genetic Epidemiol. 40, 304–314 (2016).

22. Hartwig, F. P., Davey Smith, G. & Bowden, J. Robust inference in summary data Mendelian randomization via the zero modal pleiotropy assumption. Int. J. Epidemiol. 46, 1985–1998 (2017).

23. Zhao, Q., Wang, J., Hemeni, G., Bowden, J. & Small, D. Statistical inference in two-sample summary-data Mendelian randomization using robust adjusted profile score. Annals of Statistics, 48, 1742–1769 (2020).

24. Liu, M., Jiang, Y., Wedow, R., Li, Y., Brael, D. M., Chen, F. et al. Association studies of up to 1.2 million individuals yield new insights into the genetic etiology of tobacco and alcohol use. Nat. Genet. 51, 237–244 (2019).

25. Zhou, W., Liu, G., H, R., Haycock, P., C., Aldrich, M., C., Andrew, A. S. et al. Causal relationships between body mass index, smoking and lung cancer: univariable and multivariable Mendelian randomization. Int. J. Cancer https://doi.org/10.1002/ijc.33392 (2020).

26. Mavaddat, N., Michailidou, K., Dennis, J., Lush, M., Fachal, L., Lee, A. et al. Polygenic risk scores for prediction of breast cancer and breast cancer subtypes. Am. J. Hum. Genet. 104, 21–34 (2019).

27. Bowden, J., Davey Smith, G. & Burgess, S. Mendelian randomization with invalid instruments: effect estimation and bias detection through Egger regression. Int. J. Epidemiol. 44, 512–525 (2015).

28. Bowden, J., Davey Smith, G., Haycock, P. C., Burgess, S. Consistent estimation in Mendelian randomization with some invalid instruments using a weighted median estimator. Genetic Epidemiol. 40, 304–314 (2016).

29. Sanderson, E., Davey Smith, G., Windmeijer, F. & Bowden, J. An examination of multivariable Mendelian randomization in the single-sample and two-sample summary data settings. Int. J. Epidemiol. 48, 713–727 (2019).

30. Pullit, S. L., Stoneman, C., Morris, A. P., Wood, A. R., Glastonbury, C. A., Tyrrell, J. et al. Meta-analysis of genome-wide association studies for body fat distribution in 694 649 individuals of European ancestry. Hum. Mol. Genet. 28, 166–174 (2019).

31. Lee, J. J., Wedow, R., Okbay, A., Kong, E., Maghizian, O., Zacher, M. et al. Gene discovery and polygenic prediction from a genome-wide association study of educational attainment in 1.1 million individuals. Nat. Genet. 50, 1112–1121 (2018).

32. Haycock, P. C., Burgess, S., Wade, K. H., Bowden, J., Retlon, C. & Davey Smith, G. Best (but oft-forgetten) practices: the design, analysis, and interpretation of Mendelian randomization studies. Am. J. Clin. Nutr. 103, 965–976 (2016).

33. Verbanck, M., Chen, C. Y., Neale, B. & Do, R. Detection of widespread horizontal pleiotropy in causal relationships inferred from Mendelian randomization across complex traits and diseases. Nat. Genet. 50, 669–698 (2018).

34. Larsson, S. C., Carter, P., Kar, S., Vithayathil, M., Mason, A. M., Michaelsson, K. et al. Smoking, alcohol consumption, and cancer: a Mendelian randomization study in UK Biobank and international genetic consortia participants. PLoS Med. 17, e1003178–e1003178 (2020).

35. Dossus, L., Bousquet-Ruault, M., Kaaks, R., Gram, I., Viller, A., Fervers, B. et al. Active and passive cigarette smoking and breast cancer risk: results from the EPIC cohort. Int. J. Cancer 134, 1871–1888 (2014).

36. Graham, I. T., Little, M. A., Lund, E. & Braaten, T. The fraction of breast cancer attributable to smoking: the Norwegian women and cancer study 1991–2012. Br. J. Cancer 115, 616–623 (2016).

37. Gram, I. T., Park, S. Y., Kolonel, L. N., R., Wilkens, L. L., Henderson, B. E. et al. Smoking and risk of breast cancer in a racially/ethnically diverse population of mainly women who do not drink alcohol: the MEC study. Am. J. Epidemiol. 182, 917–925 (2015).

38. Maudet, M. G., Gapturst, S. M., Sun, J., Diver, W. R., Hanan, L. M. & Thrush, M. J. Active smoking and breast cancer risk: original cohort data and meta-analysis. J. Natl Cancer Inst. 105, 515–525 (2013).
Hana A. Park1,2, Sonja Neumeyer1, Kyriaki Michailidou3,4,5, Manjeet K. Bolla6, Qin Wang5, Joe Dennis47, Dijana Plaseska-Karan130, Melissa A. Troester108, Thérèse17,55, Henrik Flyger56, Lin Fritschi57, Montserrat García-115, Marjanka K. Schmidt85,116, Rita K.77, Wei He43, Bernadette A. M. Heemskerk-Gerritsen41, Simon S. Cross41,42, Kamila Czene43, Mary81,82, Esther M.78, John L.53,54, Børge G. Nordestgaard23,24,25, Kenneth Of23x737

1Division of Cancer Epidemiology, German Cancer Research Center (DKFZ), Heidelberg, Germany; 2Faculty of Medicine, University of Heidelberg, Heidelberg, Germany; 3Biostatistics Unit, The Cyprus Institute of Neurology & Genetics, Nicosia, Cyprus; 4Cyprus School of Molecular Medicine, The Cyprus Institute of Neurology & Genetics, Nicosia, Cyprus; 5Centre for Cancer Genetic Epidemiology, Department of Public Health and Primary Care, University of Cambridge, Cambridge, UK; 6Division of Cancer Epidemiology and Genetics, National Cancer Institute, National Institutes of Health, National Institutes of Health, Department of Health and Human Services, Bethesda, MD, USA; 7Fred A. Litwin Center for Cancer Genetics, Lunenfeld-Tanenbaum Research Institute of Mount Sinai Hospital, Toronto, ON, Canada; 8Department of Molecular Genetics, University of Toronto, Toronto, ON, Canada; 9Department of Epidemiology, Genetic Epidemiology Research Institute, University of California Irvine, Irvine, CA, USA; 10N.N. Alexandrov Research Institute of Oncology and Cancer. Avaiable athttps://dietandcancerreport.org.

Update Project Expert Report 2018. Diet, nutrition, physical activity and breast cancer. Available at https://dietandcancerreport.org.
