

RESEARCH ARTICLE

Open Access



# Smoking and risk of breast cancer in the Generations Study cohort

Michael E. Jones<sup>1\*</sup>, Minouk J. Schoemaker<sup>1</sup>, Lauren B. Wright<sup>1</sup>, Alan Ashworth<sup>2,3,4,5</sup> and Anthony J. Swerdlow<sup>1,2</sup>

## Abstract

**Background:** Plausible biological reasons exist regarding why smoking could affect breast cancer risk, but epidemiological evidence is inconsistent.

**Methods:** We used serial questionnaire information from the Generations Study cohort (United Kingdom) to estimate HRs for breast cancer in relation to smoking adjusted for potentially confounding factors, including alcohol intake.

**Results:** Among 102,927 women recruited 2003–2013, with an average of 7.7 years of follow-up, 1815 developed invasive breast cancer. The HR (reference group was never smokers) was 1.14 (95% CI 1.03–1.25;  $P = 0.010$ ) for ever smokers, 1.24 (95% CI 1.08–1.43;  $P = 0.002$ ) for starting smoking at ages < 17 years, and 1.23 (1.07–1.41;  $P = 0.004$ ) for starting smoking 1–4 years after menarche. Breast cancer risk was not statistically associated with interval from initiation of smoking to first birth ( $P$ -trend = 0.97). Women with a family history of breast cancer (ever smoker vs never smoker HR 1.35; 95% CI 1.12–1.62;  $P = 0.002$ ) had a significantly larger HR in relation to ever smokers ( $P$  for interaction = 0.039) than women without (ever smoker vs never smoker HR 1.07; 95% CI 0.96–1.20;  $P = 0.22$ ). The interaction was prominent for age at starting smoking ( $P = 0.003$ ) and starting smoking relative to age at menarche ( $P = 0.0001$ ).

**Conclusions:** Smoking was associated with a modest but significantly increased risk of breast cancer, particularly among women who started smoking at adolescent or peri-menarcheal ages. The relative risk of breast cancer associated with smoking was greater for women with a family history of the disease.

**Keywords:** Smoking, Breast neoplasms, Cohort studies

## Background

The carcinogenic potential of tobacco smoke is unarguable [1, 2], and there are plausible biological reasons why smoking could affect breast cancer risk [2–5]. Authors of reviews of the association between cigarette smoking and breast cancer up to 2004 did not, however, generally find conclusive evidence for a causal relationship in humans [5–7]. Authors of more recent epidemiological analyses have reported modest raised risks with current [8–19] or former [8–15, 20] smoking, but questions remain about the extent to which this association is a consequence of confounding by alcohol use, whether risk is increased if smoking starts in adolescence or before first childbirth, and whether risk is modified by family history of breast cancer [1, 2]. We therefore examined the risk of invasive breast cancer in relation to

smoking in a large cohort study using detailed questionnaire information at recruitment and during follow-up, with adjustment for alcohol consumption and other potentially confounding factors.

## Methods

The Generations Study is a cohort study of over 113,700 women aged 16 years or older from the United Kingdom, from whom questionnaire information and informed consent were gained at recruitment since 2003 [21]. Initial recruits to the cohort were from women involved in the breast cancer charity that funded the study, as well as from women who responded to publicity about the study. Women who joined the study were asked to nominate female friends and family members, who were then contacted about joining the study. This referral method continued with subsequent recruits [21]. The first follow-up questionnaire (2½ years after recruitment) was completed by 99% of non-deceased

\* Correspondence: michael.jones@icr.ac.uk

<sup>1</sup>Division of Genetics and Epidemiology, The Institute of Cancer Research, London SW7 3RP, UK

Full list of author information is available at the end of the article



participants, a second (6 years after recruitment) by 96%, and a third (9½ years after recruitment) by 94% (of those recruited long enough ago to have entered this round of follow-up). The study was approved by the South East Multi-Centre Research Ethics Committee.

Breast and other cancers occurring in the cohort were identified from recruitment and follow-up questionnaires, spontaneous reports to the study centre, and from ‘flagging’ (*see below*) for those lost to questionnaire follow-up. Confirmation of diagnosis was obtained from cancer registries in the United Kingdom, ‘flagging’ at the National Health Service Central Registers (virtually complete registers of the populations of England and Wales, and of Scotland, to which study participants can be linked and on which deaths, cancer registrations and emigrations are ‘flagged’ and then periodically reported to authorized medical researchers), pathology reports and correspondence with patients’ general practitioners.

Information on risk factors for breast cancer was obtained from recruitment and follow-up questionnaires. In relation to smoking, women were asked if they had ‘ever smoked regularly (i.e., most days for at least 6 months); if they still smoked regularly, age started and stopped, and number of cigarettes smoked per day in different periods of their lives (during ages 16–24, 25–49, 50+ years). For analysis, we defined the period of ‘current smoking’ to include both current smokers and the year immediately after stopping, to avoid potential ‘reverse-causation’ bias from women who may have stopped smoking during the workup to a formal breast cancer diagnosis. For alcohol use, we asked women if they had been a regular drinker ‘in the sense of drinking at least one glass of alcohol per week on average; ages started and stopped, and quantity consumed in different periods of life (during ages 18–24, 25–49, 50+ years). We converted the quantity of alcohol consumed in each period of life into daily grams of alcohol. We split into three groups the women who reported current drinking (<60 g/day, ≥60 g/day, and amount unknown), and we classified women who had reported stopping drinking as former drinkers. For some women, we did not know their current drinking status during follow-up, but we knew they had consumed alcohol in the past, and these women were classified as ‘ever drinkers’. Because we had collected ages or dates at which certain events or changes in lifestyle occurred, we were able to update smoking status, alcohol use, parity, oral contraceptive (OC) use, menopausal hormone therapy (MHT) use and menopausal status at the ages these episodes occurred through the time of the second follow-up questionnaire. We updated duration of smoking for current smokers, as well as time since cessation for former smokers, in yearly increments, using smoking start and stop ages from the recruitment and second follow-up questionnaires. We

updated cigarettes smoked per day, pack-years smoked, alcohol consumption and post-menopausal body mass index (BMI) at the date of the second follow-up questionnaire.

### Statistical analysis

The present analytic cohort is based on all women who were recruited into the study during June 2003–December 2013 without prior invasive or in situ breast cancer or other malignant cancer (except non-melanoma skin cancer) or prior mastectomy. The recruitment cut-off at December 2013 was selected because at the time of analysis the second follow-up was practically complete for this group of recruits, two-thirds of the cohort had reached the third follow-up, and we had ‘flagging’ information to June 2017. Women entered risk at their date of recruitment and were censored at the earliest date of invasive breast cancer or in situ breast cancer; other malignancy (except non-melanoma skin cancer); death; most recent follow-up questionnaire (depending on date of recruitment) if completed, or the date the most recent follow-up questionnaire was due if cancer and vital status was known from ‘flagging’; or previously completed questionnaire if lost to follow-up. We censored follow-up at in situ breast cancer or other malignancy because we reasoned that if smoking is related to risk of in situ breast cancer or other malignancy, and ensuing treatments or their consequences alter risk of subsequent invasive breast cancer, including subsequent follow-up may obscure associations between smoking and invasive breast cancer.

Left-truncated and right-censored Cox proportional hazards regression [22] with attained age as the implicit time scale was used to estimate HR and 95% CI for smoking and risk of first invasive breast cancer. We adjusted for time since recruitment to cohort (0, 1–2, 3+ years); birth cohort (1908–1939, 1940–1949, 1950–1959, 1960–1969, 1970–1996); benign breast disease (yes/no); family history of breast cancer in first-degree relatives (yes/no); socio-economic score (Acorn score (<https://acorn.caci.co.uk/>) as trend, or missing indicator); age at menarche (trend, or missing indicator); age at first pregnancy (trend, or missing indicator); parity (trend, or missing indicator); duration of breastfeeding (trend, or missing indicator); current OC use during follow-up, before menopause (yes/no); alcohol consumption (trend for current drinker 1 to <60 g/day, indicator variables for never regular, current drinker ≥60 g/day, past drinker, drinker with unknown details); physical activity [log(metabolic equivalent) trend, missing indicator]; premenopausal BMI at age 20 years (trend, or missing indicator); post-menopausal BMI (trend, or missing indicator); MHT use (never used, ex-user, current oestrogen-only user, current oestrogen plus progestogen user,

current user of other types, missing indicator); menopausal status (pre- or post-menopausal); and age at menopause (trend, or missing indicator). BMI was used to create two separate variables: pre-menopausal BMI (potentially available for all women) and post-menopausal BMI (only available at post-menopausal ages). We used BMI at age 20 to represent pre-menopausal BMI. Separately, if a woman was post-menopausal at entry to the cohort we used her BMI at entry for her post-menopausal BMI (and if she was pre-menopausal at this time, post-menopausal BMI was unknown). If a woman was post-menopausal at the time of the follow-up questionnaire, we updated from this point in time her post-menopausal BMI with the value from this follow-up questionnaire. Statistical trends were evaluated using continuous values, except for duration and time since cessation of smoking, which were based on discrete time-varying, annually updated values. For trend analyses where there was an unexposed group (e.g., never smokers in analyses of smoking duration), the unexposed group was not assigned a zero magnitude but was treated as a separate categorical term, as was any missing value group. In particular we adjusted our analyses of smoking and breast cancer for alcohol using daily current alcohol consumption as a continuous measure if within the range 1 to < 60 g/day, and categorical terms for non-drinkers, for those with consumption  $\geq$  60 g/day (because we did not want a minority of women who reported very high consumption to influence unduly the trend with daily consumption), past drinkers, and those for whom details of consumption were missing, by fitting appropriate interaction terms in the Cox regression model. Heterogeneity in HRs by subtype of breast cancer defined by oestrogen receptor (ER) status or morphology was assessed using a data augmentation method [23] and Wald chi-square tests [24]. All statistical tests were two-sided, and analyses were conducted using Stata/IC version 14.0 software [25].

## Results

During 2003–2013, a recruitment questionnaire was completed by 102,940 women who had no previous invasive or in situ breast cancer or other malignancy (except non-melanoma skin cancer). At censoring date, 1.1% of women had died. Of the remainder, cancer and vital status were known for 96.5% who had completed the relevant follow-up questionnaire and for a further 2.4% from ‘flagging’ at the National Health Service Central Registers. The remaining 1.1% were lost to follow-up at an earlier date. Thirteen women (including one with breast cancer) were excluded from subsequent analyses because of self-contradictory information for parity or smoking, leaving 102,927 subjects for analysis.

Table 1 presents descriptive characteristics at recruitment of the cohort eligible for analysis. The median age

**Table 1** Characteristics of 102,927 women from the Generations Study, recruited 2003–2013

Study population	No. of subjects	%
Year of birth		
1908–1939	5917	5.8
1940–1949	21,488	20.9
1950–1959	24,735	24.0
1960–1969	23,749	23.1
1970–1996	27,038	26.3
Year of recruitment		
2003–2005	33,807	32.9
2006–2007	48,489	47.1
2008–2013	20,631	20.0
Age at recruitment, years		
16–34	23,289	22.6
35–44	23,116	22.5
45–54	24,385	23.7
55–64	23,837	23.2
65–74	7308	7.1
75–102	992	1.0
Age at menarche, years		
7–11	20,658	20.1
12–14	62,291	60.5
15–19	9353	9.1
Not known or anomalous <sup>a</sup>	10,625	10.3
Family history of breast cancer in first-degree relatives		
No	87,030	84.6
Yes	15,897	15.4
Smoking status, at entry into cohort		
Never smoked	66,013	64.1
Current smoker	8491	8.3
Former smoker	28,300	27.5
Ever smoker, status unknown at this time	123	0.1
Alcohol consumption, at entry into cohort		
Never regular	10,574	10.3
Current drinker (< 60 g/day)	63,690	61.9
Current drinker ( $\geq$ 60 g/day)	1280	1.2
Current drinker, amount unknown	4306	4.2
Former drinker	13,513	13.1
Ever drinker, status unknown at this time	9564	9.3
Total number of subjects	102,927	100.0

Total number of subjects excludes 13 women (including 1 who developed breast cancer) with self-contradictory information for parity or smoking  
<sup>a</sup>Includes 2 women reporting menarche before age 5 years, 23 at ages 20–33 years, and 23 reporting never having periods

at recruitment was 47 years (IQR 36–57). A majority of participants (64.1%) reported never smoking, but only 10.3% were never-regular consumers of alcohol. In relation to alcohol consumption, 12.5% of never smokers

were non-drinkers, in contrast to 6.4% of ever smokers. Among those who reported drinking <60 g/day, the median alcohol consumption (g/day) was 14.2 (IQR 8.7–22.1) among never smokers and 19.0 (IQR 11.9–29.2) among ever smokers. Additional file 1: Table S1 provides further descriptive characteristics of the cohort in relation to age at starting smoking, thelarche, parity, menopausal status and BMI.

During 788,361 person-years (median 6.6 years; mean 7.7 years) of follow-up, 1815 invasive breast cancers were diagnosed, of which 1813 were confirmed through national cancer registration or medical records, and the remaining 2 were self-reported with treatments that imply breast cancer. ER status data were available for 99.3%, and of these 83.7% were ER-positive. Invasive ductal carcinoma accounted for 78.8%, and lobular 16.4%, of tumours. Further descriptive characteristics of the breast cancer cases are given in Additional file 1: Table S2.

The HR for invasive breast cancer in relation to ever smoking was 1.17 (95% CI 1.07–1.29;  $P = 0.0009$ ) when adjusted only for attained age, 1.13 (95% CI 1.03–1.24;  $P = 0.012$ ) when also adjusted for alcohol consumption, and 1.14 (95% CI 1.03–1.25;  $P = 0.010$ ) when further adjusted for other potentially confounding variables (see the Methods section above and Table 2). All subsequent results are adjusted for attained age, alcohol consumption and the potentially confounding variables, unless otherwise stated.

Table 2 presents results for breast cancer overall and by ER status. The HR for ever smoking was raised for ER-positive (HR 1.12; 95% CI 1.01–1.24;  $P = 0.035$ ) and ER-negative (HR 1.25; 95% CI 0.99–1.58;  $P = 0.063$ ) breast cancer, and the difference between the HRs was not significant ( $P = 0.40$ ). Breast cancer risk increased significantly with number of cigarettes smoked per day for all breast cancer ( $P$ -trend = 0.0060) and for ER-positive tumours ( $P$ -trend = 0.023). Breast cancer risks were raised significantly after 10+ years' duration of smoking (10+ years vs never smoking  $P = 0.0004$ ). Breast cancer risks did not further rise beyond 10 years' duration, and because of this non-linear relationship, there was no significant linear trend with duration of smoking ( $P$ -trend = 0.24), nor was there significant heterogeneity in the trend by ER status. Pack-years of smoking was associated with breast cancer risk overall ( $P$ -trend = 0.0069) and ER-positive breast cancer ( $P$ -trend = 0.024) but not for ER-negative ( $P$ -trend = 0.16) tumours; there was no significant heterogeneity of the pack-years trend by ER status ( $P = 0.66$ ).

The HR within the year after smoking cessation was 2.68 (95% CI 1.60–4.46), based on 15 cases, but for reasons described in the Methods section above, this risk period was assigned for further analysis to the 'current smoker' group. On this basis, risk of breast cancer was raised in current (HR 1.12; 95% CI 0.89–1.39;  $P = 0.34$ ) and former (HR 1.14; 95% CI 1.03–1.26;  $P = 0.011$ )

smokers, although only the latter reached statistical significance; there was no significant heterogeneity by ER status. Breast cancer risks were significantly raised within the first 20 years after cessation of smoking and decreased with greater time since cessation, although the trend was not significant ( $P$ -trend = 0.071), and there was no significant heterogeneity in this trend by ER status.

There was significant variation in risk of breast cancer by age at start of smoking (Table 3) ( $P$ -heterogeneity = 0.018; not presented in Table 3). Breast cancer risk was significantly increased if smoking started at age <17 years (HR 1.24; 95% CI 1.08–1.43;  $P = 0.0023$ ) or 17–19 years (HR 1.15; 95% CI 1.01–1.31;  $P = 0.030$ ) relative to non-smokers, but not if it started at older ages. The risk was significantly increased for ER-positive subjects, only for smokers starting at ages <17 years, and no significant risk increase was noted for ER-negative breast cancer. When adjusted for pack-years, the breast cancer risk (HR) for starting smoking at age <17 years was 1.12 (95% CI 0.96–1.32;  $P = 0.14$ ), and when adjusted for duration of smoking, it was 1.16 (95% CI 0.96–1.40;  $P = 0.11$ ) (not presented in Table 3).

In our questionnaire, we asked women only about the amount they smoked per day beginning at age 16 years; therefore, we could not examine smoking intensity at younger ages. There was no significant trend in breast cancer risk, however, in relation to cigarettes smoked per day at ages 16–24 years. Relative to age at menarche, breast cancer risks were highest if smoking started at or before menarche (HR 1.40; 95% CI 0.98–1.99;  $P = 0.061$ ) or 1–4 years after (HR 1.23; 95% CI 1.07–1.41;  $P = 0.0040$ ), with a significant downward trend in breast cancer risk with increasing interval from age at menarche to age at starting smoking ( $P = 0.031$ ). A similar pattern was seen for ER-positive, but was less clear for ER-negative, breast cancer. A weaker relationship was seen with age at thelarche (e.g., 1–4 years after thelarche; HR 1.17; 95% CI 1.00–1.37;  $P = 0.056$ ). When adjusted for pack-years of smoking, the HRs for age at start of smoking 1–4 years after menarche (HR 1.12; 95% CI 0.96–1.31;  $P = 0.15$ ) or thelarche (HR 1.05; 95% CI 0.88–1.25;  $P = 0.59$ ) were attenuated (not presented in Table 3). There was a comparable attenuation after adjusting for duration of smoking. Among parous women, there was a significant trend in breast cancer risk with interval from starting smoking to birth of first child ( $P$ -trend = 0.013); for an interval of 15+ years, the HR was 1.46 (95% CI 1.18–1.81;  $P = 0.0005$ ). However, these results were not adjusted for age at first childbirth and parity (not presented in the tables), and when we adjusted (as shown in Table 3), there were no significantly raised HRs or trends for all breast cancer or by ER status.

When we analysed data by morphological type (Additional file 1: Table S3), we found significant

**Table 2** Relative risk of breast cancer in relation to smoking status, intensity, duration, dose, and cessation, by oestrogen receptor status

	Oestrogen receptor status						All breast cancer <sup>a</sup>					
	Positive			Negative			Positive			Negative		
	Cases	HR <sup>b</sup>	95% CI	P value	Cases	HR <sup>b</sup>	95% CI	P value	Cases	HR <sup>b</sup>	95% CI	P value
Ever previously smoked cigarettes <sup>c</sup>												
Never	898	1.00	Baseline		167	1.00	Baseline		1073	1.00	Baseline	
Ever	611	1.12	1.01–1.24	0.035	127	1.25	0.99–1.58	0.063	742	1.14	1.03–1.25	0.010
Ever smoking by oestrogen receptor status interaction P = 0.40												
Cigarettes per day, averaged over years when smoking <sup>d</sup>												
Never smoked	898	1.00	Baseline		167	1.00	Baseline		1073	1.00	Baseline	
1–4	216	1.07	0.92–1.25	0.36	41	1.11	0.79–1.56	0.54	258	1.08	0.94–1.24	0.29
5–9	117	1.27	1.05–1.55	0.016	17	0.97	0.59–1.59	0.89	135	1.22	1.02–1.47	0.031
10–14	68	1.37	1.07–1.76	0.013	13	1.34	0.76–2.36	0.31	82	1.37	1.09–1.72	0.0070
15+	51	1.48	1.11–1.97	0.0079	11	1.61	0.87–2.97	0.13	62	1.48	1.14–1.92	0.0031
Amount unknown	159	0.94	0.79–1.12	0.51	45	1.46	1.05–2.03	0.025	205	1.03	0.88–1.20	0.71
Trend <sup>e</sup> P = 0.023												
Trend interaction P = 0.52												
Duration of smoking, years <sup>f</sup>												
Never smoked	898	1.00	Baseline		167	1.00	Baseline		1073	1.00	Baseline	
1–9	139	0.95	0.79–1.13	0.57	37	1.30	0.91–1.85	0.15	177	1.00	0.85–1.18	0.97
10–19	187	1.22	1.04–1.43	0.015	38	1.26	0.88–1.79	0.21	225	1.21	1.05–1.41	0.009
20–29	113	1.16	0.95–1.41	0.15	28	1.59	1.06–2.38	0.025	141	1.21	1.02–1.45	0.033
30+	137	1.23	1.03–1.48	0.024	20	1.07	0.67–1.71	0.78	159	1.22	1.02–1.44	0.026
Duration unknown	35	0.97	0.69–1.36	0.86	4	0.61	0.23–1.66	0.34	40	0.93	0.68–1.28	0.66
Trend <sup>e</sup> P = 0.17												
Trend interaction P = 0.44												
Pack-years of smoking <sup>d</sup>												
Never smoked	898	1.00	Baseline		167	1.00	Baseline		1073	1.00	Baseline	
1 to < 5	156	1.13	0.95–1.34	0.15	25	0.93	0.61–1.42	0.75	182	1.10	0.94–1.29	0.25
5 to < 10	84	1.03	0.82–1.29	0.81	19	1.21	0.75–1.94	0.44	103	1.05	0.85–1.28	0.66
10 to < 20	114	1.25	1.03–1.53	0.027	24	1.42	0.92–2.18	0.11	138	1.27	1.06–1.52	0.010
20+	98	1.48	1.19–1.83	0.0003	14	1.22	0.71–2.13	0.47	114	1.45	1.19–1.77	0.0002
Amount unknown	159	0.95	0.80–1.13	0.59	45	1.47	1.06–2.05	0.022	205	1.03	0.88–1.20	0.70
Trend <sup>e</sup> P = 0.85												
Trend interaction P = 0.24												



**Table 2** Relative risk of breast cancer in relation to smoking status, intensity, duration, dose, and cessation, by oestrogen receptor status (Continued)

	Oestrogen receptor status				All breast cancer <sup>a</sup>					
	Positive		Negative		Cases		HR <sup>b</sup>		P value	
	Cases	HR <sup>b</sup>	95% CI	P value	Cases	HR <sup>b</sup>	95% CI			
	Trend <sup>d</sup> P = 0.024		Trend <sup>d</sup> P = 0.16		Trend <sup>d</sup> P = 0.0069					
	Trend interaction P = 0.66									
Current or former smoking <sup>c</sup>										
Never smoked	898	1.00	Baseline		167	1.00	Baseline	1073	1.00	Baseline
Current <sup>e</sup>	73	1.14	0.89–1.45	0.30	14	1.07	0.62–1.86	87	1.12	0.89–1.39
Former	538	1.12	1.00–1.25	0.047	113	1.27	1.00–1.62	655	1.14	1.03–1.26
	Smoking status by oestrogen receptor status interaction P = 0.58									
Time since cessation, years <sup>f</sup>										
Never smoked	898	1.00	Baseline		167	1.00	Baseline	1073	1.00	Baseline
Current <sup>e</sup> smoker	73	1.14	0.90–1.46	0.28	14	1.08	0.62–1.87	87	1.12	0.90–1.40
1–9	100	1.25	1.01–1.54	0.040	26	1.48	0.98–2.23	127	1.28	1.06–1.55
10–19	115	1.17	0.96–1.42	0.12	28	1.42	0.94–2.13	145	1.21	1.02–1.45
20–29	133	1.10	0.91–1.32	0.33	20	0.94	0.59–1.51	153	1.07	0.90–1.27
30+	170	1.05	0.88–1.24	0.61	38	1.43	0.99–2.06	209	1.10	0.94–1.28
Duration unknown	20	1.06	0.68–1.65	0.81	1	0.29	0.04–2.08	21	0.93	0.60–1.43
	Trend <sup>d</sup> P = 0.055		Trend <sup>d</sup> P = 0.85		Trend <sup>d</sup> P = 0.071					
	Trend interaction P = 0.55									

<sup>a</sup>Includes 12 breast cancers with unknown or unmeasured oestrogen receptor status

<sup>b</sup>Adjusted for attained age (Cox regression time scale); time since recruitment to cohort (0, 1–2, 3+ years); birth cohort (1908–1939, 1940–1949, 1950–1959, 1960–1969, 1970–1996); benign breast disease (yes/no); family history of breast cancer in first-degree relatives (yes/no); socio-economic score (Acorn score as trend, missing); age at menarche (trend, missing); age at first pregnancy (trend, missing); parity (trend, missing); duration of breastfeeding (trend, missing); current oral contraceptive use before menopause (yes/no); alcohol consumption (never regular, trend current drinker 1 to < 60 g/day, current drinker ≥ 60 g/day, past drinker, drinker with unknown details); physical activity (log(metabolic equivalent) trend, missing); pre-menopausal body mass index at age 20 years (trend, missing); post-menopausal body mass index (trend, missing); menopausal hormone therapy use (never used, ex-user, current oestrogen plus progestogen user, current user of other types, missing); menopausal status (pre- or post-menopausal) and age at menopause (trend, missing)

<sup>c</sup>Time updated through follow-up period

<sup>d</sup>Time updated at the point renewed information was available from follow-up questionnaire

<sup>e</sup>Trend excludes never smoker and unknown group

<sup>f</sup>Time updated in yearly steps

<sup>g</sup>Includes current smokers and < 1 year immediately after stopping

**Table 3** Relative risk of breast cancer in relation to smoking in relation to age at start of smoking, by oestrogen receptor status

	Oestrogen receptor status								All breast cancer <sup>a</sup>			
	Positive				Negative				Cases	HR <sup>b</sup>	95% CI	P value
	Cases	HR <sup>b</sup>	95% CI	P value	Cases	HR <sup>b</sup>	95% CI	P value				
Age started smoking, years <sup>c</sup>												
Never smoked	898	1.00	Baseline		167	1.00	Baseline		1073	1.00	Baseline	
< 17	218	1.25	1.08–1.46	0.0032	42	1.22	0.87–1.72	0.25	261	1.24	1.08–1.43	0.0023
17–19	249	1.13	0.98–1.30	0.093	54	1.32	0.97–1.80	0.077	304	1.15	1.01–1.31	0.030
20+	122	0.95	0.79–1.15	0.62	28	1.25	0.84–1.87	0.28	151	1.00	0.84–1.18	0.96
Age unknown	22	0.94	0.62–1.44	0.79	3	0.73	0.23–2.30	0.59	26	0.94	0.64–1.39	0.76
	Trend <sup>d</sup> P = 0.17				Trend <sup>d</sup> P = 0.74				Trend among ever smokers <sup>d</sup> P = 0.18			
	Trend interaction P = 0.30											
Cigarettes per day at ages 16–24 years (among those aged ≥ 25 years at entry)												
Never smoked	895	1.00	Baseline		166	1.00	Baseline		1069	1.00	Baseline	
1–4	77	1.10	0.87–1.39	0.42	21	1.65	1.05–2.60	0.032	99	1.19	0.97–1.46	0.10
5–9	131	1.23	1.02–1.48	0.027	18	0.92	0.57–1.50	0.75	149	1.16	0.99–1.40	0.069
10–14	170	1.11	0.94–1.31	0.22	31	1.07	0.73–1.57	0.73	203	1.10	0.95–1.29	0.21
15+	169	1.13	0.96–1.34	0.14	42	1.44	1.02–2.03	0.038	212	1.18	1.01–1.37	0.034
Started at age 25+	31	0.91	0.64–1.31	0.62	7	1.22	0.57–2.60	0.61	38	0.95	0.69–1.31	0.75
Unknown	31	0.97	0.68–1.39	0.87	7	1.32	0.62–2.79	0.47	38	1.01	0.73–1.40	0.95
	Trend <sup>d</sup> P = 0.70				Trend <sup>d</sup> P = 0.65				Trend <sup>d</sup> P = 0.78			
	Trend interaction P = 0.81											
Starting smoking relative to age at menarche, years												
Never smoked	898	1.00	Baseline		167	1.00	Baseline		1073	1.00	Baseline	
At or before menarche	32	1.60	1.12–2.28	0.010	2	0.50	0.12–2.02	0.33	34	1.40	0.98–1.99	0.061
Years started after menarche												
1–4	217	1.21	1.04–1.41	0.015	48	1.38	1.00–1.91	0.049	266	1.23	1.07–1.41	0.0040
5–9	239	1.09	0.94–1.27	0.23	54	1.35	0.99–1.84	0.060	295	1.13	0.99–1.29	0.071
10–14	36	1.03	0.74–1.44	0.86	7	1.13	0.53–2.43	0.75	43	1.04	0.76–1.41	0.82
15+	12	0.64	0.36–1.13	0.12	6	1.92	0.85–4.34	0.12	18	0.81	0.51–1.30	0.39
Interval unknown	75	1.02	0.79–1.32	0.88	10	0.74	0.38–1.41	0.35	86	0.99	0.77–1.27	0.92
	Trend <sup>d</sup> P = 0.041				Trend <sup>d</sup> P = 0.80				Trend <sup>d</sup> P = 0.031			
	Trend for interaction P = 0.42											
Starting smoking relative to age at thelarche, years												
Never smoked	898	1.00	Baseline		167	1.00	Baseline		1073	1.00	Baseline	
At or before thelarche	16	1.45	0.88–2.39	0.14	1	0.45	0.06–3.24	0.43	17	1.27	0.79–2.06	0.33
Years started after thelarche												
1–4	160	1.17	0.99–1.39	0.075	31	1.17	0.79–1.72	0.43	192	1.17	1.00–1.37	0.056
5–9	238	1.06	0.92–1.22	0.47	60	1.42	1.06–1.91	0.020	300	1.11	0.98–1.27	0.10
10–14	37	0.98	0.70–1.36	0.90	11	1.60	0.87–2.95	0.13	48	1.07	0.80–1.43	0.66
15+	18	0.94	0.59–1.50	0.79	6	1.83	0.81–4.14	0.14	24	1.06	0.71–1.59	0.77
Interval unknown	142	1.22	1.02–1.46	0.031	18	0.87	0.53–1.42	0.59	161	1.17	0.99–1.38	0.074
	Trend <sup>d</sup> P = 0.58				Trend <sup>d</sup> P = 0.35				Trend among ever smokers <sup>d</sup> P = 0.82			
	Trend for interaction P = 0.28											
Starting smoking relative to first childbirth, among parous (follow-up time only from first childbirth)												
Never smoked	772	1.00	Baseline		148	1.00	Baseline		927	1.00	Baseline	
After first childbirth	34	1.20	0.85–1.71	0.30	9	1.85	0.94–3.67	0.076	43	1.28	0.94–1.76	0.12

**Table 3** Relative risk of breast cancer in relation to smoking in relation to age at start of smoking, by oestrogen receptor status (Continued)

	Oestrogen receptor status								All breast cancer <sup>a</sup>			
	Positive				Negative				Cases	HR <sup>b</sup>	95% CI	P value
	Cases	HR <sup>b</sup>	95% CI	P value	Cases	HR <sup>b</sup>	95% CI	P value				
Interval from starting smoking to first childbirth, years												
1–4	73	1.12	0.88–1.44	0.35	15	1.28	0.74–2.21	0.38	89	1.14	0.91–1.45	0.23
5–9	177	1.06	0.90–1.25	0.51	42	1.35	0.96–1.90	0.089	220	1.10	0.94–1.28	0.22
10–14	130	1.08	0.89–1.30	0.42	22	0.89	0.57–1.40	0.62	152	1.04	0.87–1.24	0.65
15+	77	1.17	0.91–1.51	0.21	16	1.13	0.66–1.92	0.65	94	1.17	0.93–1.47	0.19
Interval unknown	22	1.07	0.70–1.64	0.75	2	0.54	0.13–2.17	0.38	25	1.02	0.69–1.52	0.92
	Trend <sup>d</sup> $P=0.72$				Trend <sup>d</sup> $P=0.33$				Trend among ever smokers <sup>d</sup> $P=0.97$			
	Trend for interaction $P=0.27$											

<sup>a</sup>Includes 21 breast cancers with unknown or unmeasured oestrogen receptor status

<sup>b</sup>Adjusted for attained age (Cox regression time scale); time since recruitment to cohort (0, 1–2, 3+ years); birth cohort (1908–1939, 1940–1949, 1950–1959, 1960–1969, 1970–1996); benign breast disease (yes/no); family history of breast cancer in first-degree relatives (yes/no); socio-economic score (Acorn score as trend, missing); age at menarche (trend, missing); age at first pregnancy (trend, missing); parity (trend, missing); duration of breastfeeding (trend, missing); current oral contraceptive use before menopause (yes/no); alcohol consumption (never regular, trend current drinker 1 to < 60 g/day, current drinker  $\geq$  60 g/day, past drinker, drinker with unknown details); physical activity [log(metabolic equivalent) trend, missing]; pre-menopausal body mass index at age 20 years (trend, missing); post-menopausal body mass index (trend, missing); menopausal hormone therapy use (never used, ex-user, current oestrogen-only user, current oestrogen plus progestogen user, current user of other types, missing); menopausal status (pre- or post-menopausal) and age at menopause (trend, missing)

<sup>c</sup>Time updated through follow-up period

<sup>d</sup>Trend excludes never smoker and unknown group

associations for ductal breast cancer similar to the results for breast cancer overall, as well as generally non-significant results for lobular breast cancer. There were no significant interactions by morphological type in the risk of breast cancer with smoking.

There was no raised risk of breast cancer with ever smoking in non-drinkers (HR 0.97; 95% CI 0.61–1.52;  $P=0.89$ ) but a significantly raised breast cancer risk in those who had ever been drinkers (HR 1.18; 95% CI 1.07–1.30;  $P=0.0010$ ), although the difference in HRs was not significant ( $P$ -interaction = 0.41) (Table 4). When further stratified by amount of alcohol consumed, the HRs for ever smoking among current drinkers remained raised. The results were similar when we examined breast cancer risk by drinking status for former smokers relative to never smokers (Additional file 1: Table S4).

We examined further potential risk factor interactions with smoking but found no significant interactions with parity ( $P=0.095$ ), although for nulliparous ever smoking women, there was a statistically significantly increased risk of breast cancer ( $P=0.012$ ) (Additional file 1: Table S5), or menopausal status ( $P=0.73$ ) (Additional file 1: Table S6). However, although the HR of pre-menopausal ever smokers was somewhat larger than for post-menopausal ever smokers, the former did not reach statistical significance ( $P=0.088$ ), whereas the latter did ( $P=0.040$ ). Nor did we find significant interactions with birth cohort ( $P=0.092$ ), BMI at age 20 years ( $P=0.55$ ) or post-menopausal ages ( $P=0.26$ ), but we did see a significant interaction with family history of breast cancer

( $P=0.038$ ). There were significant interactions between family history and age at start of smoking ( $P=0.0029$ ) and starting smoking relative to age at menarche ( $P=0.0001$ ) in relation to risk of breast cancer (Table 5). In particular, among women with a family history of breast cancer, HRs were raised if smoking started at age 20+ years (HR 1.56; 95% CI 1.17–2.10;  $P=0.0028$ ) or < 20 years (HR 1.26; 95% CI 1.02–1.56;  $P=0.029$ ), and if started 5+ years after menarche (HR 1.53; 95% CI 1.22–1.91;  $P=0.0002$ ), and we noted that these were somewhat different from the results among women without a family history of breast cancer.

## Discussion

In the Generations Study cohort, we found a significant but modestly raised risk of invasive breast cancer in ever and former smokers, in women who smoked more than five cigarettes per day, had 10+ pack-years of use, or had stopped for < 20 years. Researchers in some previous studies have reported similar associations with smoking [8–17, 20], cigarettes per day [9–11, 19], pack-years [9–13, 17–19, 26–29], and cessation [8, 12, 19, 26, 28], but not all studies have shown these associations [10, 11, 13, 15–17, 19, 20, 29, 30]. We saw significantly raised risk with 10+ years' duration of smoking, but no increasing trend beyond 10+ years. Increased risks at long durations (or significant trends) have previously been reported in some studies [8–13, 18–20, 26–28], although some classified non-smokers as smokers with zero duration [12, 20, 26, 28], and this may artefactually produce a significant trend which partly or wholly reflects the



**Table 4** Relative risk of breast cancer in relation to ever smoking, by level of alcohol consumption

Strata	Total number of cases in strata	HR for ever smoking relative to never smoking within strata of alcohol consumption <sup>a</sup>	95% CI	P value
Alcohol <sup>b</sup>				
Non-drinker	110	0.97	0.61–1.52	0.89
Ever drinker	1705	1.18	1.07–1.30	0.0010
Interaction $P^c = 0.41$				
Alcohol <sup>b</sup>				
Non-drinker	110	0.97	0.61–1.52	0.89
< 20 g/day	706	1.17	1.00–1.36	0.049
20 to < 40 g/day	356	1.17	0.95–1.44	0.13
40 to < 60 g/day	87	1.12	0.73–1.72	0.61
≥ 60 g/day	27	1.92	0.78–4.76	0.16
Current, amount unknown	180	1.47	1.09–1.98	0.011
Former drinker	295	0.98	0.78–1.25	0.89
Drinker, details missing	54	0.73	0.38–1.42	0.35
Interaction $P^c = 0.33$				

<sup>a</sup>Adjusted for attained age (Cox regression time scale); time since recruitment to cohort (0, 1–2, 3+ years); birth cohort (1908–1939, 1940–1949, 1950–1959, 1960–1969, 1970–1996); benign breast disease (yes/no); family history of breast cancer in first-degree relatives (yes/no); socio-economic score (Acorn score as trend, missing); age at menarche (trend, missing); age at first pregnancy (trend, missing); parity (trend, missing); duration of breastfeeding (trend, missing); current oral contraceptive use before menopause (yes/no); physical activity [log(metabolic equivalent) trend, missing]; pre-menopausal body mass index at age 20 years (trend, missing); post-menopausal body mass index (trend, missing); menopausal hormone therapy use (never used, ex-user, current oestrogen-only user, current oestrogen plus progestogen user, current user of other types, missing); menopausal status (pre- or post-menopausal) and age at menopause (trend, missing)

<sup>b</sup>Time updated through follow-up period

<sup>c</sup>Interaction across all categories, including missing groups

difference in risk between non-smokers and smokers (but this may not be the only reason for an association with 20+ years [long duration] of smoking).

We found that risk was significantly raised in former smokers, as previously reported [8–15, 20]. Risk was also raised with current smoking, but the number of current smokers in our cohort was small, and this result did not reach statistical significance, although researchers in some other studies have reported significantly raised risks in this group [8–19]. The raised risks for current and former smokers were similar (HR 1.12 and 1.14, respectively), and the CIs overlapped, suggesting, within our cohort, no material difference between current and former smokers in relation to breast cancer risk.

#### Breast cancer subtypes

We found significant raised risks for ER-positive and ductal breast cancer, which were the most common types in our study, but no significant heterogeneity by ER status or morphological type of the breast cancer in relation to smoking. The statistical power to examine differences by ER status or morphology was low in our cohort because of the relative uncommonness of ER-negative and non-ductal-type tumours. Some studies have tended to demonstrate stronger risks for ER-positive breast cancer [12, 16, 20, 31], but none have shown significant interactions, and the literature is inconclusive [2]. We observed larger HRs for smoking and

pre-menopausal, relative to post-menopausal, breast cancer, but the former did not reach statistical significance, and although the literature is variable, it does in general suggest a greater relative risk among pre-menopausal women [1, 2]. However, we found no evidence for a significant interaction with menopausal status, similar to other studies [8, 11, 32].

#### Confounding by alcohol

Alcohol consumption was associated with smoking and is itself a known risk factor for breast cancer [7]. We adjusted for alcohol intake, and although this reduced the strength of the association between smoking and breast cancer (from HR 1.17 to 1.14), the association remained raised and significant. There is, however, concern that statistical adjustment using self-reported alcohol consumption may not be adequate to control fully for confounding by alcohol [7], so to explore further the extent of potential confounding, we stratified by alcohol consumption (Table 4). Within each stratum of consumption (< 20 g/day, 20–40 g/day, and 40 to < 60 g/day), the difference in self-reported alcohol intake between never and ever smokers was  $\approx 1$  g/day, and we calculated that this difference in consumption would be associated with a < 1% change in relative risk of breast cancer (using the alcohol-breast cancer estimate of relative risk from a large collaborative re-analysis [7]). Within each of these strata, it would require ever smokers to

**Table 5** Relative risk of breast cancer in relation to smoking, by family history of breast cancer

	Family history of breast cancer <sup>a</sup>							
	No				Yes			
	Cases	HR <sup>b</sup>	95% CI	P value	Cases	HR <sup>b</sup>	95% CI	P value
Ever previously smoked cigarettes <sup>c</sup>								
Never	819	1.00	Baseline		239	1.00	Baseline	
Ever	554	1.07	0.96–1.20	0.22	203	1.35	1.12–1.62	0.0018
Interaction $P = 0.038$								
Age at start of smoking, years <sup>c</sup>								
Never smoker	819	1.00	Baseline		254	1.00	Baseline	
< 20	427	1.17	1.04–1.32	0.0098	138	1.26	1.02–1.56	0.029
20+	96	0.82	0.67–1.02	0.076	55	1.56	1.17–2.10	0.0028
Age unknown	16	0.75	0.46–1.23	0.26	10	1.57	0.84–2.96	0.16
Interaction $P = 0.0029$								
Trend <sup>d</sup> $P = 0.034$				Trend <sup>d</sup> $P = 0.38$				
Trend interaction $P = 0.044$								
Starting smoking relative to age at menarche, years								
Never smoker	819	1.00	Baseline		254	1.00	Baseline	
Before or < 5 years	239	1.30	1.12–1.51	0.0006	61	1.06	0.80–1.41	0.68
5+ years	244	0.97	0.84–1.13	0.70	112	1.53	1.22–1.91	0.0002
Interval unknown	56	0.83	0.62–1.12	0.22	30	1.51	1.02–2.24	0.041
Interaction $P = 0.0001$								
Trend <sup>d</sup> $P = 0.0021$				Trend <sup>d</sup> $P = 0.40$				
Trend interaction $P = 0.0079$								
Cigarettes per day age among ever smokers, averaged over years when smoking <sup>e</sup>								
Never smoker	819	1.00	Baseline		254	1.00	Baseline	
1–4	185	1.01	0.86–1.18	0.95	73	1.32	1.01–1.71	0.040
5+	210	1.28	1.09–1.49	0.0019	69	1.42	1.09–1.86	0.010
Unknown	144	0.94	0.79–1.13	0.51	61	1.31	0.99–1.74	0.059
Interaction $P = 0.13$								
Trend <sup>d</sup> $P = 0.0019$				Trend <sup>d</sup> $P = 0.81$				
Trend interaction $P = 0.18$								
Duration of smoking among ever smokers, years <sup>f</sup>								
Never smoker	819	1.00	Baseline		254	1.00	Baseline	
1–9	129	0.94	0.78–1.14	0.54	48	1.20	0.88–1.64	0.24
10+	385	1.16	1.03–1.32	0.019	140	1.39	1.13–1.71	0.0020
Duration unknown	25	0.76	0.51–1.13	0.17	15	1.51	0.90–2.55	0.12
Interaction $P = 0.094$								
Trend <sup>d</sup> $P = 0.15$				Trend <sup>d</sup> $P = 0.86$				
Trend interaction $P = 0.37$								
Pack-years of smoking among ever smokers <sup>e</sup>								
Never smoker	819	1.00	Baseline		254	1.00	Baseline	
1 to < 10	208	1.02	0.87–1.18	0.84	77	1.28	0.99–1.65	0.060
10+	187	1.30	1.11–1.53	0.0015	65	1.48	1.13–1.95	0.0051

**Table 5** Relative risk of breast cancer in relation to smoking, by family history of breast cancer (Continued)

	Family history of breast cancer <sup>a</sup>							
	No				Yes			
	Cases	HR <sup>b</sup>	95% CI	P value	Cases	HR <sup>b</sup>	95% CI	P value
Unknown	144	0.93	0.78–1.12	0.46	61	1.30	0.98–1.72	0.069
Interaction $P = 0.16$								
Trend <sup>d</sup> $P = 0.055$				Trend <sup>d</sup> $P = 0.54$				
Trend interaction $P = 0.37$								

<sup>a</sup>First-degree relatives, at recruitment

<sup>b</sup>Adjusted for attained age (Cox regression time scale); time since recruitment to cohort (0, 1–2, 3+ years); birth cohort (1908–1939, 1940–1949, 1950–1959, 1960–1969, 1970–1996); benign breast disease (yes/no); socio-economic score (Acorn score as trend, missing); age at menarche (trend, missing); age at first pregnancy (trend, missing); parity (trend, missing); duration of breastfeeding (trend, missing); current oral contraceptive use before menopause (yes/no); alcohol consumption (never regular, trend current drinker 1 to < 60 g/day, current drinker  $\geq$  60 g/day, past drinker, drinker with unknown details); physical activity [log(metabolic equivalent) trend, missing]; pre-menopausal body mass index at age 20 years (trend, missing); post-menopausal body mass index (trend, missing); menopausal hormone therapy use (never used, ex-user, current oestrogen-only user, current oestrogen plus progestogen user, current user of other types, missing); menopausal status (pre- or post-menopausal) and age at menopause (trend, missing)

<sup>c</sup>Time updated through follow-up

<sup>d</sup>Trend excludes never smoker and unknown groups

<sup>e</sup>Time updated at the point renewed information was available from follow-up questionnaire

<sup>f</sup>Time updated in yearly steps

be drinking 20 g/day more than never smokers to produce a difference of  $\approx$  15% (similar to the 12–17% we saw). This implies that the association we observed between ever smoking and breast cancer may be too large to be explained by differences in alcohol intake alone.

We saw no significant association between smoking and breast cancer risk among non-drinkers, in concordance with a collaborative re-analysis of 43 case-control and 10 cohort studies [7], the American Cancer Society's Cancer Prevention Study II cohort [16], and a subsequent pooled analysis of 14 cohort studies [8]. It is possible that there may be synergistic interaction between ever smoking and alcohol consumption, and risk of breast cancer, although only one study has shown the interaction as statistically significant [8]. There is some precedent to invoke synergism between smoking and alcohol because, for example, there is an established positive interaction between these two exposures and the aetiology of head and neck cancers [33]. However, non-drinking may occur for cultural or religious reasons, or because of underlying illness or other health issues, and in the United Kingdom at least, non-drinkers are a minority group; therefore, this potential interaction could be a reflection of a particular distribution of breast cancer risk factors among non-drinkers (and inadequate control for confounding among drinkers). Conversely, three other cohort studies demonstrated significantly raised risk among non-drinkers [18, 26, 29], although in two the increased raised risks were only in subgroups [26, 29].

### Smoking in adolescence

On the basis of epidemiological considerations and animal studies, the period from puberty to first birth may

represent a window of particular susceptibility to breast cancer [34–37]. At puberty, the breast is made up of mainly undifferentiated terminal ductal and lobular structures which animal studies show are sensitive to chemical carcinogenesis [34]. At these young ages, ionizing radiation exposure also increases risk of breast cancer [37], especially if exposure is within 6 months of menarche [38]. We found that risk of breast cancer in ever smokers was greatest if smoking started at age < 17 years or started at peri-menarcheal or, more weakly, at peri-thelarcheal ages. A number of other studies have also demonstrated raised risks if smoking started in adolescence [8–13, 16–18, 20, 26, 28, 29, 32] or around menarche [11, 16, 26]. However, when we adjusted for pack-years of smoking, the raised risks for starting smoking close to age at menarche or thelarche were somewhat attenuated, suggesting over-adjustment (because of possible correlation between age starting smoking and pack-years) or confounding by pack-years. Previous studies have not made this adjustment, so the relative importance of early initiation or pack-years of use remains unclear.

### Smoking before first childbirth

Young age at first birth and increasing parity confer long-term protection against breast cancer [34, 35], and animal models point to terminal differentiation of breast tissue at full-term pregnancy being important in this process [34–36]. Increased risks have been reported for invasive breast cancer if smoking started before first childbirth [8–11, 16, 17, 20, 26, 28, 29, 32], but we found the association was significant only if we did not adjust for age at first pregnancy. Researchers in a number of previous studies have adjusted for age at first pregnancy and still found significant associations with interval to first

birth [8, 9, 11–13, 16–18, 20, 26, 28, 29]; however, it is difficult to determine the adequacy of adjustment. For example, in a large pooled analysis of 14 cohort studies, there was a strong trend with smoking interval before first birth after adjustment for potential confounders that included age at first birth and number of live births ( $P = 0.000002$ ), whereas after stratification by age at first birth, the trends in each stratum were weaker ( $P = 0.12, 0.02,$  and  $0.28$ ) [8], which is suggestive of confounding.

### Interaction with family history

We found the association between smoking and breast cancer was significantly larger among women with a family history of the disease than in those without. Investigators in five previous studies have reported on this interaction with family history. Two studies showed no significant interaction but the researchers did not present stratified results, so we cannot determine if the direction of interaction supports or contradicts our findings [16, 19]. In three studies, researchers reported significant interactions, with one showing increased breast cancer risk with smoking only among those with a positive family history [39], whereas two demonstrated that breast cancer risk was raised only among those with no family history [15, 18]. Increased risk of breast cancer with smoking has also been seen in some [40, 41], but not all (see review [1] and a large meta-analysis [41]), studies of *BRCA1/2* carriers (or by proxy, women with three or more first-degree relatives with breast or ovarian cancer [42]). There are also reports of significant interactions with smoking and polymorphisms in carcinogen metabolism genes *NAT2* [43] and *CYP1A1* [44, 45] as well as breast cancer susceptibility single-nucleotide polymorphisms [46, 47]. Moreover, *BRCA1* and *BRCA2* proteins are involved in the repair of DNA damage, and it is therefore possible that *BRCA1/2* carriers may be more sensitive to effects of carcinogens in cigarette smoke. Thus, despite the limited and inconsistent literature, it is possible that there are gene-smoking interactions in relation to breast cancer risk (as there is, for instance, with bladder cancer [48]), and studies may benefit from being focussed on more detailed measures and timing of exposure (e.g., peri-menarcheal smoking or pack-years of use) rather than just ever/never smoking.

As in previous studies, we excluded from analysis women with prevalent breast or other malignant cancer [11–13, 15–17, 20, 28, 32] or prevalent in situ breast cancer [13] at recruitment; we restricted the analysis to invasive breast cancer [7–18, 20, 26, 28, 30]; and we adjusted for menopausal status and BMI [8, 10, 11, 13, 16, 18–20, 26, 30, 31], potential confounding variables that may also be influenced by smoking. There was little scope for bias from unascertained mortality or exits, or for erroneous reporting of breast cancer, because follow-

up for vital and breast cancer status was obtained for 99% of participants and confirmation of reported breast cancers for over 99%. Our smoking information was gained at recruitment and from follow-up questionnaires 6 years later, and we were able to update smoking status, so that women who gave up smoking were classified as former smokers from that point in time. Only a small number of other cohort studies [13, 16, 20] have been able to update smoking exposure through follow-up. One limitation of our study is that we have no direct information on passive (secondhand) smoking, and therefore our risk estimates might be underestimated if never smokers were exposed to passive smoking and if this exposure affects risk of breast cancer [49].

If our results are not due to chance, residual confounding, or unidentified bias, they suggest that certain biological mechanisms deserve further attention, such as those involving exposure at peri-menarcheal ages, and gene-environment interactions, either of which may be the direct result of chemical carcinogenesis or an indirect consequence on hormonal pathways during this susceptible period of breast development.

### Conclusions

We found that smoking was associated with a modest but significantly increased risk of breast cancer, particularly among those who started at adolescent or peri-menarcheal ages, and that the relative risk of breast cancer associated with smoking was significantly greater for women with a family history of the disease.

### Additional file

**Additional file 1: Table S1.** Further characteristics of 102,927 women from the Generations Study who were recruited 2003–2013. **Table S2.** Characteristics of incident invasive breast cancer cases arising in 102,927 women from the Generations Study. **Table S3.** Relative risk of breast cancer in relation to smoking, by morphological type. **Table S4.** Relative risk of breast cancer in relation to former smoking, by level of alcohol consumption. **Table S5.** Relative risk of breast cancer in relation to smoking, by parity. **Table S6.** Relative risk of breast cancer in relation to smoking, by menopausal status. (DOCX 72 kb)

### Abbreviations

BMI: Body mass index; ER: Oestrogen receptor; MHT: Menopausal hormone therapy; OC: Oral contraceptive

### Acknowledgements

We thank Breast Cancer Now and The Institute of Cancer Research for support of the Generations Study, as well as the study participants; the study staff; and the doctors, nurses and other health care providers and health information sources who have contributed to the study.

### Funding

This work was funded by Breast Cancer Now and The Institute of Cancer Research. We acknowledge National Health Service (NHS) funding to the Royal Marsden NHS Foundation Trust and The Institute of Cancer Research National Institute for Health Research (NIHR) Biomedical Research Centre. The funders did not participate in the design of the study, the analysis or interpretation of data, or in the writing or approval of the manuscript.

**Availability of data and materials**

The datasets generated during and/or analysed during the present study are not publicly available, owing to confidentiality reasons, but anonymised versions may be available from the corresponding author on reasonable request.

**Authors' contributions**

AJS and AA designed the study and obtained funding for the Generations Study. AJS, MEJ and MJS set up and collected data in the Generations Study. MEJ, MJS and LBW collected and prepared data for the analysis. MEJ conducted the analyses and drafted the manuscript. All authors contributed to data interpretation and preparation of the final manuscript. All authors read and approved the final manuscript.

**Ethics approval and consent to participate**

The study was approved by the South Thames Multicentre Research Ethics Committee (reference MREC 03/01/014), and participants provided informed consent.

**Consent for publication**

Not applicable.

**Competing interests**

The authors declare that they have no competing interests.

**Publisher's Note**

Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

**Author details**

<sup>1</sup>Division of Genetics and Epidemiology, The Institute of Cancer Research, London SW7 3RP, UK. <sup>2</sup>Division of Breast Cancer Research, The Institute of Cancer Research, London SW7 3RP, UK. <sup>3</sup>Breakthrough Breast Cancer Research Centre, The Institute of Cancer Research, London SW7 3RP, UK. <sup>4</sup>Division of Molecular Pathology, The Institute of Cancer Research, London SW7 3RP, UK. <sup>5</sup>Present Address: UCSF Helen Diller Family Comprehensive Cancer Center, San Francisco, CA 94158, USA.

Received: 23 August 2017 Accepted: 11 October 2017

Published online: 22 November 2017

**References**

- IARC Working Group on the Evaluation of Carcinogenic Risks to Humans. Personal habits and indoor combustions. Volume 100 E. A review of human carcinogens. IARC Monogr Eval Carcinog Risks Hum. 2012;100(Pt E):1–538.
- National Center for Chronic Disease Prevention and Health Promotion, Office on Smoking and Health, US Department of Health and Human Services. The health consequences of smoking—50 years of progress: a report of the Surgeon General. Atlanta, GA: Centers for Disease Control and Prevention; 2014.
- Morabia A. Smoking (active and passive) and breast cancer: epidemiologic evidence up to June 2001. *Environ Mol Mutagen*. 2002;39(2–3):89–95.
- Johnson KC, Miller AB, Collishaw NE, Palmer JR, Hammond SK, Salmon AG, Cantor KP, Miller MD, Boyd NF, Millar J, et al. Active smoking and secondhand smoke increase breast cancer risk: the report of the Canadian Expert Panel on Tobacco Smoke and Breast Cancer Risk (2009). *Tob Control*. 2011;20(1), e2.
- Terry PD, Rohan TE. Cigarette smoking and the risk of breast cancer in women: a review of the literature. *Cancer Epidemiol Biomarkers Prev*. 2002;11(10 Pt 1):953–71.
- IARC Working Group on the Evaluation of Carcinogenic Risks to Humans. Tobacco smoke and involuntary smoking. IARC Monogr Eval Carcinog Risks Hum. 2004;83:1–1438.
- Hamajima N, Hirose K, Tajima K, Rohan T, Calle EE, Heath Jr CW, Coates RJ, Liff JM, Talamini R, Chantarakul N, et al. Alcohol, tobacco and breast cancer – collaborative reanalysis of individual data from 53 epidemiological studies, including 58,515 women with breast cancer and 95,067 women without the disease. *Br J Cancer*. 2002;87(11):1234–45.
- Gaudet MM, Carter BD, Brinton LA, Falk RT, Gram IT, Luo J, Milne RL, Nyante SJ, Weiderpass E, Beane Freeman LE, et al. Pooled analysis of active cigarette smoking and invasive breast cancer risk in 14 cohort studies. *Int J Epidemiol*. 2017;46(3):881–93.
- Bjerkaas E, Parajuli R, Weiderpass E, Engeland A, Maskarinec G, Selmer R, Gram IT. Smoking duration before first childbirth: an emerging risk factor for breast cancer? Results from 302,865 Norwegian women. *Cancer Causes Control*. 2013;24(7):1347–56.
- Catsburg C, Miller AB, Rohan TE. Active cigarette smoking and risk of breast cancer. *Int J Cancer*. 2015;136(9):2204–9.
- Dossus L, Boutron-Ruault MC, Kaaks R, Gram IT, Vilier A, Fervers B, Manjer J, Tjonneland A, Olsen A, Overvad K, et al. Active and passive cigarette smoking and breast cancer risk: results from the EPIC cohort. *Int J Cancer*. 2014;134(8):1871–88.
- Luo J, Margolis KL, Wactawski-Wende J, Horn K, Messina C, Stefanick ML, Tindle HA, Tong E, Rohan TE. Association of active and passive smoking with risk of breast cancer among postmenopausal women: a prospective cohort study. *BMJ*. 2011;342:d1016.
- Xue F, Willett WC, Rosner BA, Hankinson SE, Michels KB. Cigarette smoking and the incidence of breast cancer. *Arch Intern Med*. 2011;171(2):125–33.
- Macacu A, Autier P, Boniol M, Boyle P. Active and passive smoking and risk of breast cancer: a meta-analysis. *Breast Cancer Res Treat*. 2015;154(2):213–24.
- Nyante SJ, Gierach GL, Dallal CM, Freedman ND, Park Y, Danforth KN, Hollenbeck AR, Brinton LA. Cigarette smoking and postmenopausal breast cancer risk in a prospective cohort. *Br J Cancer*. 2014;110(9):2339–47.
- Gaudet MM, Gapstur SM, Sun J, Diver WR, Hannan LM, Thun MJ. Active smoking and breast cancer risk: original cohort data and meta-analysis. *J Natl Cancer Inst*. 2013;105(8):515–25.
- Gram IT, Little MA, Lund E, Braaten T. The fraction of breast cancer attributable to smoking: the Norwegian Women and Cancer Study 1991–2012. *Br J Cancer*. 2016;115(5):616–23.
- Reynolds P, Hurley S, Goldberg DE, Anton-Culver H, Bernstein L, Deapen D, Horn-Ross PL, Peel D, Pinder R, Ross RK, et al. Active smoking, household passive smoking, and breast cancer: evidence from the California Teachers Study. *J Natl Cancer Inst*. 2004;96(1):29–37.
- Cui Y, Miller AB, Rohan TE. Cigarette smoking and breast cancer risk: update of a prospective cohort study. *Breast Cancer Res Treat*. 2006;100(3):293–9.
- Al-Delaimy WK, Cho E, Chen WY, Colditz G, Willet WC. A prospective study of smoking and risk of breast cancer in young adult women. *Cancer Epidemiol Biomarkers Prev*. 2004;13(3):398–404.
- Swerdlow AJ, Jones ME, Schoemaker MJ, Hemming J, Thomas D, Williamson J, Ashworth A. The Breakthrough Generations Study: design of a long-term UK cohort study to investigate breast cancer aetiology. *Br J Cancer*. 2011;105(7):911–7.
- Cox DR. Regression models and life-tables. *J R Stat Soc Series B Stat Methodol*. 1972;34(2):187–220.
- Lunn M, McNeil D. Applying Cox regression to competing risks. *Biometrics*. 1995;51(2):524–32.
- Buse A. The likelihood ratio, Wald, and Lagrange multiplier tests: an expository note. *Am Stat*. 1982;36(3):153–7.
- StataCorp. Stata Statistical Software: release 14. College Station, TX: StataCorp LP; 2015.
- Gram IT, Braaten T, Terry PD, Sasco AJ, Adami HO, Lund E, Weiderpass E. Breast cancer risk among women who start smoking as teenagers. *Cancer Epidemiol Biomarkers Prev*. 2005;14(1):61–6.
- Li CI, Malone KE, Daling JR. The relationship between various measures of cigarette smoking and risk of breast cancer among older women 65–79 years of age (United States). *Cancer Causes Control*. 2005;16(8):975–85.
- Luo J, Horn K, Ockene JK, Simon MS, Stefanick ML, Tong E, Margolis KL. Interaction between smoking and obesity and the risk of developing breast cancer among postmenopausal women: the Women's Health Initiative Observational Study. *Am J Epidemiol*. 2011;174(8):919–28.
- Gram IT, Park SY, Kolonel LN, Maskarinec G, Wilkens LR, Henderson BE, Le Marchand L. 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*. 2015;182(11):917–25.
- Catsburg C, Kirsh VA, Soskolne CL, Kreiger N, Rohan TE. Active cigarette smoking and the risk of breast cancer: a cohort study. *Cancer Epidemiol*. 2014;38(4):376–81.
- London SJ, Colditz GA, Stampfer MJ, Willett WC, Rosner BA, Speizer FE. Prospective study of smoking and the risk of breast cancer. *J Natl Cancer Inst*. 1989;81(21):1625–31.
- Egan KM, Stampfer MJ, Hunter D, Hankinson S, Rosner BA, Holmes M, Willett WC, Colditz GA; Nurses' Health Study. Active and passive smoking in breast cancer: prospective results from the Nurses' Health Study. *Epidemiology*. 2002;13(2):138–45.



33. Boffetta P, Hashibe M. Alcohol and cancer. *Lancet Oncol.* 2006;7(2):149–56.
34. Russo J, Mailo D, Hu YF, Balogh G, Sheriff F, Russo IH. Breast differentiation and its implication in cancer prevention. *Clin Cancer Res.* 2005;11(2 Pt 2):931s–6.
35. Colditz GA, Bohlke K, Berkey CS. Breast cancer risk accumulation starts early: prevention must also. *Breast Cancer Res Treat.* 2014;145(3):567–79.
36. Colditz GA, Frazier AL. Models of breast cancer show that risk is set by events of early life: prevention efforts must shift focus. *Cancer Epidemiol Biomarkers Prev.* 1995;4(5):567–71.
37. Carmichael A, Sami AS, Dixon JM. Breast cancer risk among the survivors of atomic bomb and patients exposed to therapeutic ionising radiation. *Eur J Surg Oncol.* 2003;29(5):475–9.
38. Cooke R, Jones ME, Cunningham D, Falk SJ, Gilson D, Hancock BW, Harris SJ, Horwich A, Hoskin PJ, Illidge T, et al. Breast cancer risk following Hodgkin lymphoma radiotherapy in relation to menstrual and reproductive factors. *Br J Cancer.* 2013;108(11):2399–406.
39. Suzuki T, Matsuo K, Wakai K, Hiraki A, Hirose K, Sato S, Ueda R, Tajima K. Effect of familial history and smoking on common cancer risks in Japan. *Cancer.* 2007;109(10):2116–23.
40. Breast Cancer Family Registry; Kathleen Cuninghams Consortium for Research into Familial Breast Cancer (Australasia); Ontario Cancer Genetics Network (Canada). 1. Smoking and risk of breast cancer in carriers of mutations in *BRCA1* or *BRCA2* aged less than 50 years. *Breast Cancer Res Treat.* 2008;109:67–75.
41. Friebel TM, Domchek SM, Rebbeck TR. Modifiers of cancer risk in *BRCA1* and *BRCA2* mutation carriers: systematic review and meta-analysis. *J Natl Cancer Inst.* 2014;106(6):dju091.
42. Couch FJ, Cerhan JR, Vierkant RA, Grabrick DM, Therneau TM, Pankratz VS, Hartmann LC, Olson JE, Vachon CM, Sellers TA. Cigarette smoking increases risk for breast cancer in high-risk breast cancer families. *Cancer Epidemiol Biomarkers Prev.* 2001;10(4):327–32.
43. Ambrosone CB, Kropp S, Yang J, Yao S, Shields PG, Chang-Claude J. Cigarette smoking, *N*-acetyltransferase 2 genotypes, and breast cancer risk: pooled analysis and meta-analysis. *Cancer Epidemiol Biomarkers Prev.* 2008;17(1):15–26.
44. Ishibe N, Hankinson SE, Colditz GA, Spiegelman D, Willett WC, Speizer FE, Kelsey KT, Hunter DJ. Cigarette smoking, cytochrome *P450 1A1* polymorphisms, and breast cancer risk in the Nurses' Health Study. *Cancer Res.* 1998;58(4):667–71.
45. Saintot M, Malaveille C, Hautefeuille A, Gerber M. Interactions between genetic polymorphism of cytochrome *P450-1B1*, sulfotransferase *1A1*, catechol-*O*-methyltransferase and tobacco exposure in breast cancer risk. *Int J Cancer.* 2003;107(4):652–7.
46. Barndahl M, Rudolph A, Hopper JL, Southey MC, Broeks A, Fasching PA, Beckmann MW, Gago-Dominguez M, Castelao JE, Guénel P, et al. Gene–environment interactions involving functional variants: results from the Breast Cancer Association Consortium. *Int J Cancer.* 2017;141(9):1830–40.
47. Rudolph A, Milne RL, Truong T, Knight JA, Seibold P, Flesch-Janys D, Behrens S, Eilber U, Bolla MK, Wang Q, et al. Investigation of gene–environment interactions between 47 newly identified breast cancer susceptibility loci and environmental risk factors. *Int J Cancer.* 2015;136(6):E685–96.
48. García-Closas M, Malats N, Silverman D, Dosemeci M, Kogevinas M, Hein DW, Tardón A, Serra C, Carrato A, García-Closas R, et al. *NAT2* slow acetylation, *GSTM1* null genotype, and risk of bladder cancer: results from the Spanish Bladder Cancer Study and meta-analyses. *Lancet.* 2005;366(9486):649–59.
49. Johnson KC. Accumulating evidence on passive and active smoking and breast cancer risk. *Int J Cancer.* 2005;117(4):619–28.

Submit your next manuscript to BioMed Central and we will help you at every step:

- We accept pre-submission inquiries
- Our selector tool helps you to find the most relevant journal
- We provide round the clock customer support
- Convenient online submission
- Thorough peer review
- Inclusion in PubMed and all major indexing services
- Maximum visibility for your research

Submit your manuscript at  
[www.biomedcentral.com/submit](http://www.biomedcentral.com/submit)

