FULL PAPER

BJC

British Journal of Cancer (2014) 110, 2339–2347 | doi: 10.1038/bjc.2014.132

Keywords: smoking; menarche; family history of breast cancer; breast cancer risk; cohort study

Cigarette smoking and postmenopausal breast cancer risk in a prospective cohort

S J Nyante*[,1], G L Gierach[1], C M Dallal[2], N D Freedman[1], Y Park[1], K N Danforth[3], A R Hollenbeck[4] and L A Brinton[1]

[1]*Division of Cancer Epidemiology and Genetics, National Cancer Institute, 9609 Medical Center Drive, Bethesda, MD 20892, USA;* [2]*Department of Epidemiology and Biostatistics, University of Maryland School of Public Health, 2234 SPH Building no. 255, College Park, MD 20742, USA;* [3]*Department of Research and Evaluation, Kaiser Permanente Southern California, 100 South Los Robles Avenue, 2nd Floor, Pasadena, CA 91101, USA and* [4]*Strategic Issues Research, AARP Research Center, 601 E Street NW, Washington, DC 20049, USA*

There is a growing body of evidence to suggest a relationship between cigarette smoking and increased breast cancer risk. Components of cigarette smoke have been detected in fluid from breast ducts of smokers (Petrakis *et al*, 1978; Hill and Wynder, 1979), and smoking has been associated with a higher level of DNA adducts (Perera *et al*, 1995; Li *et al*, 1996), higher prevalence of *TP53* mutations and p53 overexpression in breast tumours (Gammon *et al*, 1999; Conway *et al*, 2002), providing a plausible hypothesis for a biological link between smoking and risk. Yet, epidemiologic evidence has been mixed. The majority of studies have used a case–control design and reported a mixture of positive, null, and inverse associations (Terry and Rohan, 2002). But as noted by Terry and Rohan (Terry and Rohan, 2002), case–control analyses of smoking may be subject to differential recall bias; prospective assessments of smoking avoid this bias and therefore may be more appropriate. While early prospective studies reported a mixture of positive (Hiatt and Fireman, 1986) and null (Hiatt *et al*, 1988; Schatzkin *et al*, 1989; Engeland *et al*, 1996; Nordlund *et al*, 1997) associations, more recent cohort studies have reported positive associations, with relative risks ranging from 1.08 to 1.70 for current smoking and 1.00 to 1.34 for former smoking (Manjer *et al*, 2001; Al-Delaimy *et al*, 2004; Reynolds *et al*, 2004; Gram *et al*, 2005; Hanaoka *et al*, 2005; Olson *et al*, 2005; Cui *et al*, 2006; Ha *et al*, 2007; Xue *et al*, 2011; Luo *et al*, 2011b; Gaudet *et al*, 2013).

Reasons for the variation in association among prospective studies are unclear, but potential causes include confounding or

*Correspondence: Dr SJ Nyante; E-mail: sarah.nyante@nih.gov

Received 23 October 2013; revised 3 February 2014; accepted 10 February 2014; published online 18 March 2014

effect measure modification. Most recent studies have adjusted for potential confounders in regression models, but it is possible that residual confounding could affect the association between smoking and risk (Hamajima *et al*, 2002). Alcohol use, in particular, has been identified as a potential source of residual confounding due to the positive correlation between alcohol use and smoking (Moore *et al*, 2005) and the association between alcohol and increased breast cancer risk (Hamajima *et al*, 2002; Dumitrescu and Shields, 2005). In a recent study, Gaudet *et al* (2013) presented associations stratified by alcohol use status (never, former, current) but did not include information on the amount of alcohol consumed. As such, it is unclear whether associations among current and former drinkers are subject to residual confounding by the amount of alcohol consumed.

Limited prospective data also suggest that there may be differences in smoking-associated risk according to body mass index (BMI) and family history of breast cancer. In one study, smoking was associated with increased risk in non-obese women while there was no association among obese women (Luo *et al*, 2011a). In another, smoking-associated risk was stronger among women without a family history of breast cancer when compared with women with a family history (Reynolds *et al*, 2004). However, there was no modification by BMI or family history in other studies (Gram *et al*, 2005; Cui *et al*, 2006; Gaudet *et al*, 2013). There are also data to suggest that the smoking association with risk is elevated among current and former alcohol drinkers but not among non-drinkers (London *et al*, 1989; Gaudet *et al*, 2013), but these differences were not statistically significant and other studies have found no modification by alcohol use (Gram *et al*, 2005; Cui *et al*, 2006). Analyses involving menopausal status (Reynolds *et al*, 2004; Gram *et al*, 2005; Cui *et al*, 2006), menopausal hormone therapy (MHT) use (Cui *et al*, 2006; Gaudet *et al*, 2013), age (Gram *et al*, 2005), age at menarche (Cui *et al*, 2006), parity (Gram *et al*, 2005; Cui *et al*, 2006), age at first birth (Gram *et al*, 2005), and benign breast disease (Cui *et al*, 2006) suggest that these factors do not modify the association between smoking and risk, but these factors have been examined in a limited number of studies.

To clarify our understanding of the relationship between smoking and breast cancer risk, we examined the association using data from the National Institutes of Health (NIH)-AARP Diet and Health Study, a large prospective cohort study that has previously demonstrated breast cancer risk relationships for a variety of factors that could potentially confound or modify associations with cigarette smoking (Ahn *et al*, 2007; Brinton *et al*, 2008; Lew *et al*, 2009; Nyante *et al*, 2013). Additionally, we examined whether smoking was related to the risk of specific subgroups of breast cancer, defined by disease stage at diagnosis, histology, hormone receptor expression, and tumour grade.

MATERIALS AND METHODS

Study population. The NIH-AARP Diet and Health Study has been described previously (Schatzkin *et al*, 2001). Briefly, in 1995–1996 a questionnaire regarding health and nutrition was sent to members of AARP (formerly American Association of Retired Persons) who were 50–71 years old and living in California, Florida, Louisiana, New Jersey, North Carolina, Pennsylvania, Atlanta, Georgia, and Detroit, Michigan. The study was approved by the Special Studies Institutional Review Board of the National Cancer Institute. Completion of the questionnaire implied informed consent.

Of 617 119 questionnaires returned, 567 169 were completed satisfactorily (Schatzkin *et al*, 2001). For this analysis, we excluded responses that were duplicates ($N=179$) or completed by proxy ($N=15\,760$). We also excluded respondents who were male ($N=325\,172$), had prevalent cancer ($N=23\,961$), moved from the study area or died before the questionnaire was returned ($N=592$), were premenopausal or had unknown menopausal status ($N=9420$), withdrew from the study ($N=9$), or did not provide smoking information ($N=5926$), resulting in 186 150 postmenopausal women in the analytic cohort (Figure 1).

Risk factors. At enrolment, participants were asked whether they had smoked $\geqslant 100$ cigarettes during their lifetime, whether they currently smoked, when they stopped smoking, and the number of cigarettes they usually smoked per day. Smoking status was defined as follows: participants who had not smoked $\geqslant 100$ cigarettes were classified as never smokers; participants who had smoked $\geqslant 100$ cigarettes and either currently smoked or quit within 1 year of enrolment were classified as current smokers; and participants who had smoked $\geqslant 100$ cigarettes but stopped smoking >1 year prior to study enrolment were classified as former smokers.

Type of menopause was determined from the reason menstrual periods stopped, whether the ovaries had been removed, and age at last menstrual period. MHT use was determined from questions regarding how long women used 'replacement hormones' and whether they were currently using hormones. Body mass index was calculated as weight (kilograms) divided by squared height (metres). Frequency of vigorous physical activity was defined by how often women participated in exercise, sports, or carrying heavy loads that increased sweating, breathing, or heart rate and lasted $\geqslant 20$ min. Alcohol use (grams per day) was estimated from the frequency and amount of beer, wine, and liquor consumed (Lew *et al*, 2009). Family history of breast cancer was based on history in a first-degree female relative. Other covariates included age at menarche, age at first birth, previous breast biopsy, education level, and race/ethnicity. Data were missing for $<5\%$ of participants for all variables.

Cohort follow-up. Participants were followed for changes of address using data from the US Postal Service, other address update services, and participant updates. Vital status was

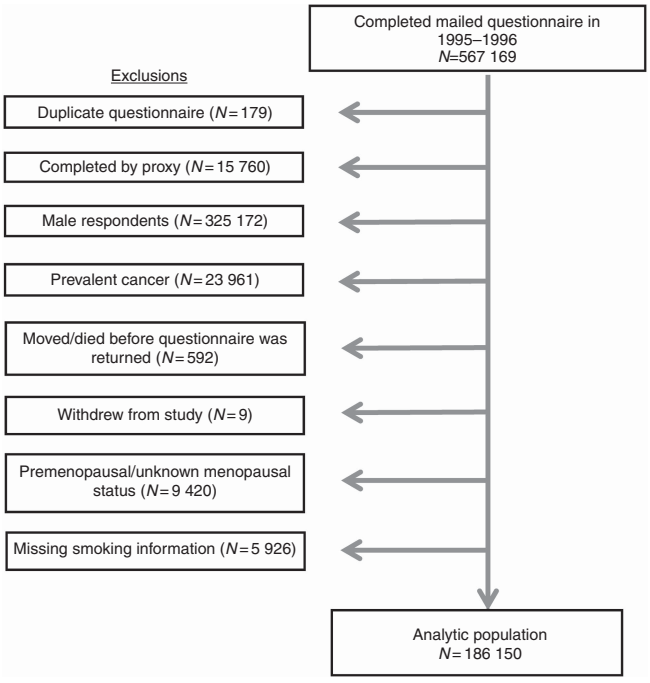

Figure 1. The relationship between smoking and breast cancer risk was examined in 186 150 female, postmenopausal participants in the NIH-AARP Diet and Health Study cohort. Exclusion criteria are shown in the figure.

determined through periodic linkage to the Social Security Administration Death Master File, responses to study mailings, National Death Index searches, and cancer registry linkages. Follow-up time was calculated from the date the questionnaire was returned through the earliest of the following: breast cancer diagnosis, movement out of the registry ascertainment area, death, or 31 December 2006. 12 790 (7%) women moved from the registry ascertainment area prior to the end of follow-up; these women were younger and more educated than the rest of the cohort, but had similar distributions of smoking status and other risk factors.

Case ascertainment. Cancer diagnoses were obtained from state cancer registries in the eight study areas plus the adjacent states of Texas, Arizona, and Nevada (Michaud *et al*, 2005). A validation study comparing registry findings with self-reports and medical records estimated that linkage identified 90% of incident cancers (Michaud *et al*, 2005). Cases were defined as primary invasive breast cancers diagnosed after study enrolment. Histology, defined using the International Classification of Diseases for Oncology (Fritz *et al*, 2000), tumour grade, hormone receptor expression, and disease stage, defined by the Surveillance, Epidemiology, and End Results (SEER) Program summary staging definitions (Young *et al*, 2001), were obtained from cancer registry records. Disease stage at diagnosis was classified as localised if the disease was confined to the breast tissue and breast fat; regional spread if the disease had spread to adjacent tissues by direct extension and/or spread to the regional lymph nodes; or distant metastases if the disease had spread to non-adjacent organ sites. Hormone receptor data were reported by California, Louisiana, New Jersey, North Carolina, Arizona, Nevada, Georgia, and Michigan registries; these states ascertained 64% of the study's breast cancer cases.

Statistical analysis. Time-dependent interactions between each variable and age during follow-up were tested for significance to establish that hazards were proportional over time (all $P > 0.05$). We estimated multivariable-adjusted hazard ratios (HR) and 95% confidence intervals (CI) for the association between smoking status, number of cigarettes smoked, and time since quitting smoking and breast cancer risk using Cox proportional hazards regression in SAS v9.2 (SAS, Cary, NC, USA); age was the underlying timescale. Models were adjusted for age at enrolment (continuous), race/ethnicity (white non-Hispanic, other), education level (high school or less, vocational/some college, college graduate, postgraduate), age at menarche ($\leqslant$12, 13–14, $\geqslant$15 years), age at first birth (nulliparous, $<$20, 20–24, 25–29, $\geqslant$30 years), type of menopause (natural menopause at $<$45, 45–49, 50–54, $\geqslant$55 years, bilateral oophorectomy, other surgery, medical/unknown age at natural menopause), MHT use (never, former, current), BMI in quartiles ($<$22.9, 22.9–25.7, 25.8–29.6, $>$29.6 kg m^{-2}), alcohol consumption (0, $>$0–5, $>$5–10, $>$10–20, $>$20–35, $>$35 g per day), frequency of vigorous physical activity (never/rarely, 1–3 times per month, 1–2 times per week, 3–4 times per week, $\geqslant$5 times per week), previous breast biopsy (no, yes), and family history of breast cancer (no, yes). To further evaluate confounding by the amount of alcohol consumed, we also estimated smoking associations within strata of alcohol consumption. Formal tests of linear trend for number of cigarettes and time since quitting were conducted by modelling exposure categories as ordinal variables among smokers only.

Effect measure modification by BMI ($<$30, $\geqslant$30 kg m^{-2}), family history of breast cancer (no, yes), alcohol consumption ($\leqslant$5 g per day, $>$5 g per day), type of menopause (natural at $<$50 years, natural at $\geqslant$50 years, bilateral oophorectomy, other surgical or medical menopause), MHT use (never, former, current), age at menarche ($\leqslant$12, 13–14, $\geqslant$15 years), age at first birth (nulliparous, $<$25, $\geqslant$25 years), previous breast biopsy (yes, no), and age at baseline ($<$55, 55–59, 60–64, $\geqslant$65 years) was evaluated using the same categorisations used in previous studies that reported interaction with these factors (Vatten and Kvinnsland, 1990; Ursin *et al*, 2005; Magnusson *et al*, 2007; Luo *et al*, 2011a) or categorisations associated with breast cancer risk. We also evaluated alternative categorisations for BMI and MHT using categorisations used in prior studies (BMI: $\leqslant$24, $>$24 kg m^{-2}; MHT: never, ever) (Vatten and Kvinnsland, 1990; Ursin *et al*, 2005). We constructed risk estimates for smoking status stratified by these factors and additionally compared multivariable-adjusted models with and without multiplicative interaction terms using the likelihood ratio test. Tests were two-sided and P-values <0.05 were considered statistically significant. The expected joint HR under the multiplicative null was calculated by multiplying the HR associated with only smoking (unexposed to the potential modifier) and the HR associated with only being exposed to the potential modifier (unexposed to smoking).

Associations between smoking and risk were estimated for specific breast cancer types, defined by disease stage, histology, oestrogen and progesterone receptor (ER, PR) expression, and tumour grade. Tests of statistical heterogeneity among tumour types were performed using case-only logistic regression. Risk of ER/PR-defined tumours was evaluated using data only from participants that lived in states where registries collected hormone receptor data.

To evaluate whether the presence of smoking-related health conditions affected the relationship with risk, we examined models that excluded women with self-reported poor health or emphysema and compared those results to the main analysis. We also estimated associations for the first and second halves of the follow-up period to address the possibility that changes in health characteristics or behaviours during the follow-up period affected risk estimates. For the first half of follow-up, person-time was truncated 5 years after enrolment and cases occurring after that time were censored. For the second half of follow-up, only person-time and incident cases occurring $>$5 years after enrolment were analysed.

A subsequent questionnaire mailed in 2004–2006 obtained additional information, including the age when participants started smoking. Of 186 150 women in the baseline cohort, 90 950 (49%) provided additional smoking information in 2004–2006. We attempted to explore whether age at smoking initiation was associated with risk in this sub-cohort of women, but found that the association between smoking status and risk in the subgroup was null, potentially due to limited follow-up, lack of power, or selection bias. Therefore, age at smoking initiation was not included in further analyses.

RESULTS

A total of 186 150 women were followed for a mean of 9.6 years (standard deviation—2.5); 17% were current smokers, 38% were former smokers, and 45% had never smoked (Supplementary Table 1). A total of 7481 breast cancers were diagnosed during follow-up. Breast cancer risk was elevated among current and former smokers as compared with never smokers (current, HR 1.19, 95% CI 1.10–1.28; former, HR 1.07, 95% CI 1.01–1.13; Table 1). There was no clear dose–response relationship between number of cigarettes smoked per day and risk among current or former smokers, nor was there a dose–response trend with time since quitting. Risk for former smokers remained slightly elevated when compared with never smokers, even for those who quit $\geqslant$10 years prior to baseline (HR 1.04, 95% CI 0.98–1.11).

To explore whether associations between smoking status and risk might be due to confounding by alcohol consumption, we estimated associations within strata of alcohol consumption and found that the association remained elevated across strata (Supplementary Table 2). Notably, current smoking was suggestive

Table 1. Cigarette smoking and postmenopausal breast cancer risk in the NIH-AARP Diet and Health Study, 1995–2006

	Person-years	Cases	HR (95% CI)[a]
Smoking status			
Never	827 507	3301	1.00 (reference)
Former	669 087	2912	1.07 (1.01–1.13)
Current	290 375	1268	1.19 (1.10–1.28)
Cigarettes smoked per day			
Never	827 507	3301	1.00 (reference)
Former smokers			
1–10	258 936	1105	1.07 (0.99–1.15)
11–20	203 078	830	1.00 (0.92–1.08)
21–30	104 912	498	1.18 (1.07–1.31)
31–40	60 464	295	1.19 (1.05–1.35)
⩾41	41 697	184	1.05 (0.89–1.23)
P-trend			0.10
Current smokers			
1–10	90 146	394	1.19 (1.06–1.34)
11–20	123 196	556	1.22 (1.11–1.35)
21–30	51 707	210	1.06 (0.91–1.23)
31–40	19 946	88	1.12 (0.88–1.41)
⩾41	5379	20	0.92 (0.57–1.49)
P-trend			0.15
Years since quit smoking[b]			
Never	827 507	3301	1.00 (reference)
⩾10	472 406	2058	1.04 (0.98–1.11)
5–9	120 876	543	1.19 (1.08–1.31)
1–4	75 805	311	1.09 (0.96–1.24)
P-trend			0.09

Abbreviations: CI = confidence interval; HR = hazard ratio.
[a]Adjusted for age at study entry, education level, race/ethnicity, age at menarche, nulliparity/age at first birth, type of and age at menopause, menopausal hormone therapy use, grams of alcohol per day, BMI, frequency of vigorous physical activity, previous breast biopsy, and family history of breast cancer.
[b]Estimated among former and never smokers.

of increased risk even among women who did not drink alcohol (HR 1.14, 95% CI 0.98–1.32).

We next examined how the association between smoking status and risk varied across strata of other risk factors (Figure 2). The association between current smoking and risk was stronger among women with BMI <30, without a family history of breast cancer, who reached menarche at ⩾15 years old, who drank ⩾5 g alcohol per day, who experienced natural menopause, who were not currently using menopausal hormones, and who were <55 years old at baseline. These differences were more pronounced for current rather than former smoking. With respect to alcohol consumption, smoking associations with risk did not differ when we separated the stratum of women who drank ⩽5 g per day into 0 g and >0 to 5 g per day (Supplementary Table 2). Smoking-associated risks did not differ notably by age at first birth or previous breast biopsy (Figure 2). The results did not change when we used alternative categorisations for BMI (⩽24, >24 kg m^{-2}) or MHT use (never, ever) used in previous studies (Supplementary Table 3).

Tests of multiplicative interaction with smoking status were significant only for family history of breast cancer (*P*-interaction = 0.03) and age at menarche (*P*-interaction = 0.03). When we examined their joint associations with smoking, the increased risk due to current smoking among women with a family history (HR 1.49, CI 1.27–1.75) was lower than the expected joint association under the multiplicative null (expected HR = 1.93) (Table 2). The joint HR for former smokers was 1.60 (CI 1.45, 1.77), similar to the expected HR of 1.68. With regards to the joint association between smoking and age at menarche, current smokers who reached menarche at ⩾15 years had 1.21 (95% CI 0.99–1.47) times the risk of never smokers who reached menarche at ⩽12 years, an association that was qualitatively different from the expected joint HR of 0.88 (Table 3).

There were no strong differences when we examined risk of specific breast cancer types. Although current and former smoking were associated more strongly with distant metastases (current HR 2.05, CI 1.30–3.23; former HR 1.56, CI 1.06–2.30) than localised (current HR 1.21, CI 1.09–1.34; former HR 1.07, CI 0.99–1.16) or regional spread (current HR 1.18, CI 0.99–1.41; former HR 1.10, CI 0.96–1.26), these differences were not significantly different (*P*-heterogeneity = 0.22; Table 4). Associations with current, but not former, smoking were slightly stronger for lobular rather than ductal or mixed ductal-lobular tumours (*P*-heterogeneity = 0.55) and for ER+/PR− rather than ER+/PR+ or ER−/PR− tumours (*P*-heterogeneity = 0.40). There were no differences in smoking-associated risk by tumour grade (*P*-heterogeneity = 0.90).

Associations between smoking status and risk were similar after excluding women who had emphysema or poor health, and also for the first and second halves of follow-up (Supplementary Table 4).

DISCUSSION

In the NIH-AARP Diet and Health study, current and former smokers had greater risk of breast cancer when compared with never smokers. Other cohort studies have shown similar associations between smoking and risk (Al-Delaimy *et al*, 2004; Reynolds *et al*, 2004; Gram *et al*, 2005; Hanaoka *et al*, 2005; Olson *et al*, 2005; Cui *et al*, 2006; Ha *et al*, 2007; Xue *et al*, 2011; Luo *et al*, 2011b; Gaudet *et al*, 2013), and a recent meta-analysis of prospective studies estimated combined hazard ratios of 1.12 for current smoking and 1.09 for former smoking (Gaudet *et al*, 2013). In our analysis, there was no dose-response with number of cigarettes smoked daily. Some (Reynolds *et al*, 2004; Gram *et al*, 2005; Cui *et al*, 2006; Xue *et al*, 2011) but not all (Al-Delaimy *et al*, 2004; Luo *et al*, 2011b; Gaudet *et al*, 2013) prior prospective studies have observed a positive trend between number of cigarettes and risk. While it is possible that the lack of a dose–response relationship in the NIH-AARP cohort could be due to the lack of a true association between smoking and breast cancer risk, this is unlikely given the consistency of associations for current and former smoking with those from previous prospective studies, including those that observed dose–response relationships with number of cigarettes smoked. Another possibility is that dose-response was not observed due to the small magnitude of the overall smoking association. Additionally, misclassification of the number of cigarettes smoked may have affected the results. Misclassification could have arisen due to improper recall of number of cigarettes smoked, the use of a single number to approximate dose when the number of cigarettes smoked daily may have varied over decades, or variation over time in cigarette brands, formulations, filters, and manufacturing processes that changed the dose of carcinogens ingested per day even if the same number of cigarettes was smoked per day. However, it is difficult to tell whether any of these reasons affected the associations we observed. The lack of association between time since quitting smoking and risk was similar to most previous prospective studies (Reynolds *et al*, 2004; Gram *et al*, 2005; Cui *et al*, 2006; Xue *et al*, 2011; Gaudet *et al*, 2013), except for one (Luo *et al*, 2011b) that observed lower risk as time since quitting increased.

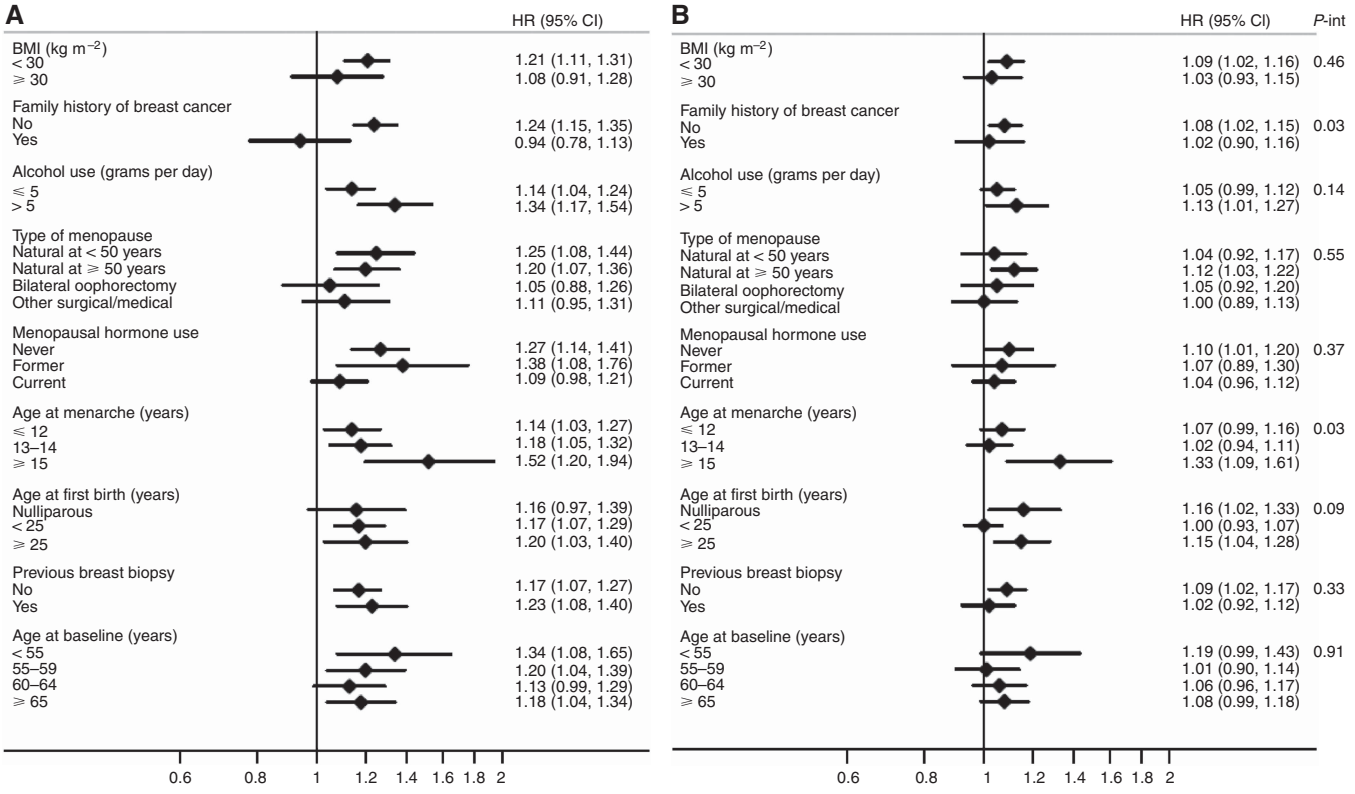

Figure 2. Association between smoking and postmenopausal breast cancer risk, stratified by breast cancer risk factors. The association between current (shown in **A**) and former (**B**) smoking and breast cancer risk, in comparison to never smoking, is shown stratified by known breast cancer risk factors. Associations were estimated using data from 186 150 postmenopausal women in the NIH-AARP Diet and Health Study. The *P*-value for interaction (*P*-int, at far right) was estimated using the likelihood ratio test by comparing models containing a multiplicative interaction term between smoking status (never, current, former) and the risk factor of interest with the models with no interaction term.

Table 2. Joint associations[a] of smoking and family history of breast cancer with postmenopausal breast cancer risk

	Family history of breast cancer	
	No	**Yes**
Smoking status		
Never		
HR (95% CI)[b]	1.00 (reference)	1.56 (1.42–1.71)
Cases	2551	617
Person-years	689 876	103 053
Former		
HR (95% CI)[b]	1.08 (1.01–1.14)	1.60 (1.45–1.77)
Cases	2289	506
Person-years	558 410	83 133
Current		
HR (95% CI)[b]	1.24 (1.14–1.34)	1.49 (1.27–1.75)
Cases	1023	174
Person-years	242 761	33 628

Abbreviations: CI = confidence interval; HR = hazard ratio.
[a]*P*-interaction = 0.03.
[b]Adjusted for age at study entry, education level, race/ethnicity, age at menarche, nulliparity/age at first birth, age at and type of menopause, menopausal hormone therapy use, grams of alcohol consumed per day, BMI, frequency of vigorous physical activity, and previous breast biopsy.

Table 3. Joint associations[a] of smoking and age at menarche with postmenopausal breast cancer risk

	Age at menarche (years)		
	⩽12	**13–14**	**⩾15**
Smoking status			
Never			
HR (95% CI)[b]	1.00 (reference)	0.98 (0.91–1.06)	0.77 (0.66–0.89)
Cases	1665	1374	251
Person-years	401 836	344 288	78 524
Former			
HR (95% CI)[b]	1.07 (0.99–1.15)	1.00 (0.92–1.08)	1.04 (0.91–1.20)
Cases	1456	1198	255
Person-years	326 005	280 890	60 495
Current			
HR (95% CI)[b]	1.14 (1.03–1.27)	1.15 (1.03–1.28)	1.21 (0.99–1.47)
Cases	620	515	131
Person-years	142 563	118 010	29 113

Abbreviations: CI = confidence interval; HR = hazard ratio.
[a]*P*-interaction = 0.03.
[b]Adjusted for age at study entry, education level, race/ethnicity, nulliparity/age at first birth, age at and type of menopause, menopausal hormone therapy use, grams of alcohol consumed per day, BMI, frequency of vigorous physical activity, previous breast biopsy, and family history of breast cancer.

Table 4. Smoking status and risk of specific breast cancer types in the NIH-AARP Diet and Health Study

	Smoking status	Cases (*N*)	HR (95% CI)[a]	*P*-heterogeneity
Disease stage at diagnosis				
Localised	Never	1571	1.00 (reference)	0.22
	Former	1404	1.07 (0.99–1.16)	
	Current	625	1.21 (1.09–1.34)	
Regional spread	Never	563	1.00 (reference)	
	Former	496	1.10 (0.96–1.26)	
	Current	217	1.18 (0.99–1.41)	
Distant metastases	Never	55	1.00 (reference)	
	Former	67	1.56 (1.06–2.30)	
	Current	42	2.05 (1.30–3.23)	
Histology				
Ductal	Never	2272	1.00 (reference)	0.55
	Former	2147	1.09 (1.02–1.17)	
	Current	757	1.17 (1.07–1.28)	
Lobular	Never	357	1.00 (reference)	
	Former	332	1.05 (0.89–1.24)	
	Current	130	1.31 (1.05–1.62)	
Mixed ductal-lobular	Never	280	1.00 (reference)	
	Former	252	1.01 (0.84–1.21)	
	Current	88	1.03 (0.79–1.35)	
Hormone receptor expression[b,c]				
ER+/PR+	Never	1148	1.00 (reference)	0.40
	Former	995	1.04 (0.95–1.14)	
	Current	369	1.02 (0.89–1.17)	
ER+/PR-	Never	222	1.00 (reference)	
	Former	212	1.13 (0.92–1.39)	
	Current	109	1.37 (1.05–1.80)	
ER−/PR−	Never	260	1.00 (reference)	
	Former	219	1.08 (0.88–1.31)	
	Current	93	1.05 (0.81–1.38)	
Tumour grade				
Low	Never	675	1.00 (reference)	0.90
	Former	675	1.12 (0.99–1.25)	
	Current	236	1.19 (1.02–1.40)	
Intermediate	Never	1277	1.00 (reference)	
	Former	1151	1.03 (0.95–1.13)	
	Current	408	1.14 (1.01–1.28)	
High	Never	868	1.00 (reference)	
	Former	809	1.10 (0.99–1.22)	
	Current	293	1.18 (1.03–1.36)	

Abbreviations: CI = confidence interval; HR = hazard ratio.
[a]Adjusted for age at study entry, education level, race/ethnicity, age at menarche, nulliparity/age at first birth, type of and age at menopause, menopausal hormone therapy use, grams of alcohol per day, BMI, frequency of vigorous physical activity, previous breast biopsy, and family history of breast cancer.
[b]Estimated only among women residing in areas that reported hormone receptor data (California, Louisiana, New Jersey, North Carolina, Arizona, Nevada, Atlanta, Georgia, and Detroit, Michigan).
[c]ER−/PR+ association not estimated due to low case numbers.

The association between smoking and risk is weak in comparison to other breast cancer risk factors, and previous investigations (Hamajima *et al*, 2002) have raised the suggestion that associations between smoking and risk may be due to confounding, in particular, due to alcohol consumption. Alcohol is a potentially strong confounder owing to its strong relationships with smoking (Moore *et al*, 2005) and breast cancer risk (Hamajima *et al*, 2002; Dumitrescu and Shields, 2005). In the NIH-AARP cohort, alcohol use was associated with risk in a dose-response manner, such that women who drank >35 g alcohol per day had 35% greater risk than non-drinkers (Lew *et al*, 2009). Gaudet *et al* (2013) showed that the association between smoking and risk was unchanged after adjusting for recency of alcohol use (never, former, current use), but such broad categories leave potential for residual confounding by alcohol dose. In our analysis, the association between smoking and risk was still observed after adjusting for the amount of alcohol consumed. Further, we estimated associations within strata of alcohol use, with the hypothesis that if the association with smoking was due to confounding by increasing alcohol dose, associations within strata

of alcohol use would be null. We found that risk associated with current smoking was elevated across all strata of alcohol use, but risk associated with former smoking was more variable. Together with the results from the study by Gaudet *et al* (2013), these results suggest that the association between smoking and risk is likely not due to confounding by alcohol. However, studies with sufficient data to simultaneously address issues of recency and dose of both smoking and alcohol use are needed to fully address this concern, particularly as it relates to former smoking and risk.

We evaluated interactions between smoking and a number of risk factors; many of the weaker variations in association may have been due to chance. Smoking-associated risks differed according to family history of breast cancer, where increased risks were only observed among women without a family history. This was consistent with the data from Reynolds *et al* (2004), who reported 39% increased risk for current smoking among postmenopausal women without a family history but no association among women with a family history. Two other studies did not find interaction between smoking and family history (Cui *et al*, 2006; Gaudet *et al*, 2013), but stratified estimates were not shown; therefore, it is unknown whether the estimates were consistent with our findings. In women at high risk of a specific type of familial breast cancer, *BRCA1/BRCA2* mutation carriers, current smoking was not associated with increased risk (Ghadirian *et al*, 2004; Nkondjock *et al*, 2006; Ginsburg *et al*, 2009), which is consistent with our finding of no smoking-associated risk among women with a family history, although smoking was associated with increased risk among mutation carriers in one study (Breast Cancer Family Registry *et al*, 2008). The potential mechanism that would result in smoking-associated risk only among those without a family history of breast cancer is unknown. It could be hypothesised that smoking and familial breast cancer contribute to risk through the same pathway, and if that pathway is activated by factors related to family history there cannot be further activation by smoking. It is also possible that the differences we observed were due to chance. Greater exploration of the role of genetic factors in smoking-associated breast carcinogenesis is needed to understand possible effect modification by family history.

In this cohort, increased risk due to smoking was also higher among women with late ages at menarche, but there are scant data from other studies to support this result. Of the two previous prospective studies that examined smoking risks stratified by age at menarche, one (Gram *et al*, 2005) showed similar results—smoking was associated with increased risk among women who reached menarche at $\geqslant 13$ years but not women with menarche at <13 years. However, this was based on a small number of cases ($N=478$) and may have been due to chance. Another study found no differences in smoking-associated risk by age at menarche (Cui *et al*, 2006).

Biologically, interaction between smoking and reproductive development is plausible. Pubertal breast development generally begins prior to menarche and marks a time of increased epithelial cell proliferation (Lanigan *et al*, 2007). As such, this may be a critical time for susceptibility to carcinogens. Other investigations point towards the importance of the timing of smoking initiation in relation to reproductive development. A recent meta-analysis reported positive associations between young age at smoking initiation and risk and smoking before first birth and risk (Gaudet *et al*, 2013). Studies have also reported that risk was elevated among women who started smoking prior to menarche (Gram *et al*, 2005; Gaudet *et al*, 2013), and that risk increased with the number of years of smoking between menarche and first birth (Xue *et al*, 2011).We were unable to evaluate the relationship between age at smoking initiation and risk in this cohort, or the effect of smoking initiation in relation to menarche or first birth. The relative importance of each reproductive milestone in relation to smoking and risk remains to be determined, but this and other studies to date indicate that smoking during these time periods may contribute to breast carcinogenesis. In both instances where we detected interaction, the increased risks were among subgroups that were at relatively lower risk (i.e., women without a family history of breast cancer and women with later age at menarche). It may be possible that increased risks due to smoking are more likely to be detected among women at low risk; however, this pattern was not consistent for all breast cancer risk factors in the stratified analyses.

There was little evidence to suggest that smoking was associated with risk of specific types of breast cancer, although risks were somewhat stronger for women with distant metastases, ER+/PR− tumours, and lobular breast tumours. Previous studies have been inconsistent, with several reporting stronger associations between smoking and ER+ as compared with ER− tumours (Al-Delaimy *et al*, 2004; Luo *et al*, 2011b; Gaudet *et al*, 2013) and one reporting a strong positive association between smoking and ER− but not ER+ tumours (Manjer *et al*, 2001). Studies that examined risk using combined ER/PR status found elevated risk of ER+/PR+ (Luo *et al*, 2011b) and ER−/PR− (Manjer *et al*, 2001) tumours, but not ER+/PR− tumours. Differences by disease stage and histology have been similarly inconsistent (Manjer *et al*, 2001; Luo *et al*, 2011b; Gaudet *et al*, 2013). Heterogeneity according to other tumour characteristics, such as molecular subtype, may exist. Studies with an emphasis on tissue collection are needed to investigate this possibility.

Our study was limited by the fact that smoking status and other covariates were assessed at baseline and do not account for changes during the follow-up period. Associations between smoking status and risk were similar when we compared the first 5 years of follow-up with the second 5 years, suggesting that changes that may have occurred during follow-up did not affect relative risks. In addition, information regarding passive smoking was not collected. Others have found that inclusion of passive smokers with never smokers yields a similar, if slightly attenuated, association (Egan *et al*, 2002; Reynolds *et al*, 2004; Luo *et al*, 2011b). Thus, if passive smoking affected this analysis it is likely that the association between active smoking and risk is stronger than what we reported here.

Smoking duration was not assessed on the baseline questionnaire and we were unable to evaluate whether breast cancer risk increased linearly with number of years of smoking, as has been reported by some studies (Gram *et al*, 2005; Olson *et al*, 2005; Cui *et al*, 2006; Luo *et al*, 2011b) but not others (Reynolds *et al*, 2004; Gaudet *et al*, 2013). Among previous studies that have evaluated both number of cigarettes (dose) and smoking duration, results for these exposures have been inconsistent: dose and duration trends were similar in four studies (Gram *et al*, 2005; Cui *et al*, 2006; Xue *et al*, 2011; Gaudet *et al*, 2013) and dissimilar in three (Al-Delaimy *et al*, 2004; Reynolds *et al*, 2004; Luo *et al*, 2011b). Thus, we cannot infer from the lack of association with number of cigarettes smoked that there is also no association with duration in the NIH-AARP population. Hormone receptor status was only available for a subset of participants; thus, these analyses were less precise than those using the full cohort. Finally, we were unable to examine age at smoking initiation, which may be an important predictor of smoking-associated risk. Further examination of when women begin smoking, in concert with the timing of menarche and first birth, may provide insight as to the mechanism by which smoking is associated with breast cancer risk.

This study also had several strengths. Smoking was assessed prospectively, reducing the potential for recall bias. We used both adjustment and stratification to evaluate potential confounding by alcohol dose and found that neither approach was able to explain the smoking association with risk. This analysis included >7000 cases, more than most previous prospective studies of smoking and breast cancer risk, allowing for precise estimation of stratified associations and associations between smoking and specific breast

cancer types. Finally, we conducted several sensitivity analyses to address potential biases due to women with poor health, respiratory diseases linked to smoking, or changes in risk factor status over the follow-up period. We found no evidence that these factors influenced the results.

CONCLUSION

Smoking status was positively associated with breast cancer risk in this cohort, but there was no dose–response trend with the number of cigarettes smoked per day or time since quitting. Risks were modified by family history of breast cancer and age at menarche, but additional research evaluating effects of smoking on the genome and on breast tissue during different stages of reproductive development are needed to better understand these findings.

ACKNOWLEDGEMENTS

This research was supported by the Intramural Research Program of the National Cancer Institute at the National Institutes of Health. Cancer incidence data from the Atlanta metropolitan area were collected by the Georgia Center for Cancer Statistics, Department of Epidemiology, Rollins School of Public Health, Emory University. Cancer incidence data from California were collected by the California Department of Health Services, Cancer Surveillance Section. Cancer incidence data from the Detroit metropolitan area were collected by the Michigan Cancer Surveillance Program, Community Health Administration, State of Michigan. The Florida cancer incidence data used in this report were collected by the Florida Cancer Data System (FCDC) under contract with the Florida Department of Health (FDOH).Cancer incidence data from Louisiana were collected by the Louisiana Tumour Registry, Louisiana State University Medical Center in New Orleans. Cancer incidence data from New Jersey were collected by the New Jersey State Cancer Registry, Cancer Epidemiology Services, New Jersey State Department of Health and Senior Services. Cancer incidence data from North Carolina were collected by the North Carolina Central Cancer Registry. Cancer incidence data from Pennsylvania were supplied by the Division of Health Statistics and Research, Pennsylvania Department of Health, Harrisburg, Pennsylvania. Cancer incidence data from Arizona were collected by the Arizona Cancer Registry, Division of Public Health Services, Arizona Department of Health Services. Cancer incidence data from Texas were collected by the Texas Cancer Registry, Cancer Epidemiology and Surveillance Branch, Texas Department of State Health Services. We are indebted to the participants in the NIH-AARP Diet and Health Study for their outstanding cooperation. We also thank Sigurd Hermansen and Kerry Grace Morrissey from Westat for study outcomes ascertainment and management and Leslie Carroll at Information Management Services for data support and analysis.

CONFLICT OF INTEREST

The authors declare no conflict of interest.

DISCLAIMER

The views expressed herein are solely those of the authors and do not necessarily reflect those of the FCDC or FDOH. The Pennsylvania Department of Health specifically disclaims responsibility for any analyses, interpretations or conclusions.

REFERENCES

Ahn J, Schatzkin A, Lacey Jr. JV, Albanes D, Ballard-Barbash R, Adams KF, Kipnis V, Mouw T, Hollenbeck AR, Leitzmann MF (2007) Adiposity, adult weight change, and postmenopausal breast cancer risk. *Arch Intern Med* **167**(19): 2091–2102.

Al-Delaimy WK, Cho E, Chen WY, Colditz G, Willet WC (2004) A prospective study of smoking and risk of breast cancer in young adult women. *Cancer Epidemiol Biomarkers Prev* **13**(3): 398–404.

Breast Cancer Family Registry, Kathleen Cuningham Consortium for Research into Familial Breast Cancer (Australasia), Ontario Cancer Genetics Network (Canada) (2008) Smoking and risk of breast cancer in carriers of mutations in BRCA1 or BRCA2 aged less than 50 years. *Breast Cancer Res Treat* **109**(1): 67–75.

Brinton LA, Richesson D, Leitzmann MF, Gierach GL, Schatzkin A, Mouw T, Hollenbeck AR, Lacey JV Jr (2008) Menopausal hormone therapy and breast cancer risk in the NIH-AARP Diet and Health Study Cohort. *Cancer Epidemiol Biomarkers Prev* **17**(11): 3150–3160.

Conway K, Edmiston SN, Cui L, Drouin SS, Pang J, He M, Tse CK, Geradts J, Dressler L, Liu ET, Millikan R, Newman B (2002) Prevalence and spectrum of p53 mutations associated with smoking in breast cancer. *Cancer Res* **62**(7): 1987–1995.

Cui Y, Miller AB, Rohan TE (2006) Cigarette smoking and breast cancer risk: update of a prospective cohort study. *Breast Cancer Res Treat* **100**(3): 293–299.

Dumitrescu RG, Shields PG (2005) The etiology of alcohol-induced breast cancer. *Alcohol* **35**(3): 213–225.

Egan KM, Stampfer MJ, Hunter D, Hankinson S, Rosner BA, Holmes M, Willett WC, Colditz GA (2002) Active and passive smoking in breast cancer: prospective results from the Nurses' Health Study. *Epidemiology* **13**(2): 138–145.

Engeland A, Andersen A, Haldorsen T, Tretli S (1996) Smoking habits and risk of cancers other than lung cancer: 28 years' follow-up of 26,000 Norwegian men and women. *Cancer Causes Control* **7**(5): 497–506.

Fritz A, Percy C, Jack A, Shanmugaratnam K, Sobin L, Parkin D, Whelan S (2000) *International Classification of Diseases for Oncology*. (ICD-O-3) 3rd edn. World Health Organization: Geneva, Switzerland.

Gammon MD, Hibshoosh H, Terry MB, Bose S, Schoenberg JB, Brinton LA, Bernstein JL, Thompson WD (1999) Cigarette smoking and other risk factors in relation to p53 expression in breast cancer among young women. *Cancer Epidemiol Biomarkers Prev* **8**(3): 255–263.

Gaudet MM, Gapstur SM, Sun J, Diver WR, Hannan LM, Thun MJ (2013) Active smoking and breast cancer risk: original cohort data and meta-analysis. *J Natl Cancer Inst* **105**(8): 515–525.

Ghadirian P, Lubinski J, Lynch H, Neuhausen SL, Weber B, Isaacs C, Baruch RG, Randall S, Ainsworth P, Friedman E, Horsman D, Tonin P, Foulkes WD, Tung N, Sun P, Narod SA (2004) Smoking and the risk of breast cancer among carriers of BRCA mutations. *Int J Cancer* **110**(3): 413–416.

Ginsburg O, Ghadirian P, Lubinski J, Cybulski C, Lynch H, Neuhausen S, Kim-Sing C, Robson M, Domchek S, Isaacs C, Klijn J, Armel S, Foulkes WD, Tung N, Moller P, Sun P, Narod SA (2009) Smoking and the risk of breast cancer in BRCA1 and BRCA2 carriers: an update. *Breast Cancer Res Treat* **114**(1): 127–135.

Gram IT, Braaten T, Terry PD, Sasco AJ, Adami HO, Lund E, Weiderpass E (2005) Breast cancer risk among women who start smoking as teenagers. *Cancer Epidemiol Biomarkers Prev* **14**(1): 61–66.

Ha M, Mabuchi K, Sigurdson AJ, Freedman DM, Linet MS, Doody MM, Hauptmann M (2007) Smoking cigarettes before first childbirth and risk of breast cancer. *Am J Epidemiol* **166**(1): 55–61.

Hamajima N, Hirose K, Tajima K, Rohan T, Calle EE, Heath Jr CW, Coates RJ, Liff JM, Talamini R, Chantarakul N, Koetsawang S, Rachawat D, Morabia A, Schuman L, Stewart W, Szklo M, Bain C, Schofield F, Siskind V, Band P, Coldman AJ, Gallagher RP, Hislop TG, Yang P, Kolonel LM, Nomura AM, Hu J, Johnson KC, Mao Y, De Sanjose S, Lee N, Marchbanks P, Ory HW, Peterson HB, Wilson HG, Wingo PA, Ebeling K, Kunde D, Nishan P, Hopper JL, Colditz G, Gajalanski V, Martin N, Pardthaisong T, Silpisornkosol S, Theetranont C, Boosiri B, Chutivongse S, Jimakorn P, Virutamasen P, Wongsrichanalai C, Ewertz M, Adami HO, Bergkvist L, Magnusson C, Persson I, Chang-Claude J, Paul C, Skegg DC, Spears GF, Boyle P, Evstifeeva T, Daling JR, Hutchinson WB, Malone K, Noonan EA, Stanford JL,

Thomas DB, Weiss NS, White E, Andrieu N, Bremond A, Clavel F, Gairard B, Lansac J, Piana L, Renaud R, Izquierdo A, Viladiu P, Cuevas HR, Ontiveros P, Palet A, Salazar SB, Aristizabel N, Cuadros A, Tryggvadottir L, Tulinius H, Bachelot A, Le MG, Peto J, Franceschi S, Lubin F, Modan B, Ron E, Wax Y, Friedman GD, Hiatt RA, Levi F, Bishop T, Kosmelj K, Primic-Zakelj M, Ravnihar B, Stare J, Beeson WL, Fraser G, Bullbrook RD, Cuzick J, Duffy SW, Fentiman IS, Hayward JL, Wang DY, McMichael AJ, McPherson K, Hanson RL, Leske MC, Mahoney MC, Nasca PC, Varma AO, Weinstein AL, Moller TR, Olsson H, Ranstam J, Goldbohm RA, van den Brandt PA, Apelo RA, Baens J, de la Cruz JR, Javier B, Lacaya LB, Ngelangel CA, La Vecchia C, Negri E, Marubini E, Ferraroni M, Gerber M, Richardson S, Segala C, Gatei D, Kenya P, Kungu A, Mati JG, Brinton LA, Hoover R, Schairer C, Spirtas R, Lee HP, Rookus MA, van Leeuwen FE, Schoenberg JA, McCredie M, Gammon MD, Clarke EA, Jones L, Neil A, Vessey M, Yeates D, Appleby P, Banks E, Beral V, Bull D, Crossley B, Goodill A, Green J, Hermon C, Key T, Langston N, Lewis C, Reeves G, Collins R, Doll R, Peto R, Mabuchi K, Preston D, Hannaford P, Kay C, Rosero-Bixby L, Gao YT, Jin F, Yuan JM, Wei HY, Yun T, Zhiheng C, Berry G, Cooper Booth J, Jelihovsky T, MacLennan R, Shearman R, Wang QS, Baines CJ, Miller AB, Wall C, Lund E, Stalsberg H, Shu XO, Zheng W, Katsouyanni K, Trichopoulou A, Trichopoulos D, Dabancens A, Martinez L, Molina R, Salas O, Alexander FE, Anderson K, Folsom AR, Hulka BS, Bernstein L, Enger S, Haile RW, Paganini-Hill A, Pike MC, Ross RK, Ursin G, Yu MC, Longnecker MP, Newcomb P, Kalache A, Farley TM, Holck S, Meirik O (2002) 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* **87**(11): 1234–1245.

Hanaoka T, Yamamoto S, Sobue T, Sasaki S, Tsugane S (2005) Active and passive smoking and breast cancer risk in middle-aged Japanese women. *Int J Cancer* **114**(2): 317–322.

Hiatt RA, Fireman BH (1986) Smoking, menopause, and breast cancer. *J Natl Cancer Inst* **76**(5): 833–838.

Hiatt RA, Klatsky AL, Armstrong MA (1988) Alcohol consumption and the risk of breast cancer in a prepaid health plan. *Cancer Res* **48**(8): 2284–2287.

Hill P, Wynder EL (1979) Nicotine and cotinine in breast fluid. *Cancer Lett* **6**(4-5): 251–254.

Lanigan F, O'Connor D, Martin F, Gallagher WM (2007) Molecular links between mammary gland development and breast cancer. *Cell Mol Life Sci* **64**(24): 3159–3184.

Lew JQ, Freedman ND, Leitzmann MF, Brinton LA, Hoover RN, Hollenbeck AR, Schatzkin A, Park Y (2009) Alcohol and risk of breast cancer by histologic type and hormone receptor status in postmenopausal women: the NIH-AARP Diet and Health Study. *Am J Epidemiol* **170**(3): 308–317.

Li D, Wang M, Dhingra K, Hittelman WN (1996) Aromatic DNA adducts in adjacent tissues of breast cancer patients: clues to breast cancer etiology. *Cancer Res* **56**(2): 287–293.

London SJ, Colditz GA, Stampfer MJ, Willett WC, Rosner BA, Speizer FE (1989) Prospective study of smoking and the risk of breast cancer. *J Natl Cancer Inst* **81**(21): 1625–1631.

Luo J, Horn K, Ockene JK, Simon MS, Stefanick ML, Tong E, Margolis KL (2011a) 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* **174**(8): 919–928.

Luo J, Margolis KL, Wactawski-Wende J, Horn K, Messina C, Stefanick ML, Tindle HA, Tong E, Rohan TE (2011b) Association of active and passive smoking with risk of breast cancer among postmenopausal women: a prospective cohort study. *BMJ* **342**: d1016.

Magnusson C, Wedren S, Rosenberg LU (2007) Cigarette smoking and breast cancer risk: a population-based study in Sweden. *Br J Cancer* **97**(9): 1287–1290.

Manjer J, Malina J, Berglund G, Bondeson L, Garne JP, Janzon L (2001) Smoking associated with hormone receptor negative breast cancer. *Int J Cancer* **91**(4): 580–584.

Michaud D, Midthune D, Hermansen S, Leitzmann M, Harlan L, Kipnis V, Schatzkin A (2005) Comparison of cancer registry case ascertainment with SEER estimates and self-reporting in a subset of the NIH-AARP Diet and Health Study. *J Registry Manage* **32**(2): 70–75.

Moore AA, Gould R, Reuben DB, Greendale GA, Carter MK, Zhou K, Karlamangla A (2005) Longitudinal patterns and predictors of alcohol consumption in the United States. *Am J Public Health* **95**(3): 458–465.

Nkondjock A, Robidoux A, Paredes Y, Narod SA, Ghadirian P (2006) Diet, lifestyle and BRCA-related breast cancer risk among French-Canadians. *Breast Cancer Res Treat* **98**(3): 285–294.

Nordlund LA, Carstensen JM, Pershagen G (1997) Cancer incidence in female smokers: a 26-year follow-up. *Int J Cancer* **73**(5): 625–628.

Nyante SJ, Dallal CM, Gierach GL, Park Y, Hollenbeck AR, Brinton LA (2013) Risk factors for specific histopathological types of postmenopausal breast cancer in the NIH-AARP Diet and Health Study. *Am J Epidemiol* **178**(3): 359–371.

Olson JE, Vachon CM, Vierkant RA, Sweeney C, Limburg PJ, Cerhan JR, Sellers TA (2005) Prepregnancy exposure to cigarette smoking and subsequent risk of postmenopausal breast cancer. *Mayo Clin Proc* **80**(11): 1423–1428.

Perera FP, Estabrook A, Hewer A, Channing K, Rundle A, Mooney LA, Whyatt R, Phillips DH (1995) Carcinogen-DNA adducts in human breast tissue. *Cancer Epidemiol Biomarkers Prev* **4**(3): 233–238.

Petrakis NL, Gruenke LD, Beelen TC, Castagnoli Jr N, Craig JC (1978) Nicotine in breast fluid of nonlactating women. *Science* **199**(4326): 303–305.

Reynolds P, Hurley S, Goldberg DE, Anton-Culver H, Bernstein L, Deapen D, Horn-Ross PL, Peel D, Pinder R, Ross RK, West D, Wright WE, Ziogas A (2004) Active smoking, household passive smoking, and breast cancer: evidence from the California Teachers Study. *J Natl Cancer Inst* **96**(1): 29–37.

Schatzkin A, Carter CL, Green SB, Kreger BE, Splansky GL, Anderson KM, Helsel WE, Kannel WB (1989) Is alcohol consumption related to breast cancer? Results from the Framingham Heart Study. *J Natl Cancer Inst* **81**(1): 31–35.

Schatzkin A, Subar AF, Thompson FE, Harlan LC, Tangrea J, Hollenbeck AR, Hurwitz PE, Coyle L, Schussler N, Michaud DS, Freedman LS, Brown CC, Midthune D, Kipnis V (2001) Design and serendipity in establishing a large cohort with wide dietary intake distributions: the National Institutes of Health-American Association of Retired Persons Diet and Health Study. *Am J Epidemiol* **154**(12): 1119–1125.

Terry PD, Rohan TE (2002) Cigarette smoking and the risk of breast cancer in women: a review of the literature. *Cancer Epidemiol Biomarkers Prev* **11**(10 Pt 1): 953–971.

Ursin G, Bernstein L, Lord SJ, Karim R, Deapen D, Press MF, Daling JR, Norman SA, Liff JM, Marchbanks PA, Folger SG, Simon MS, Strom BL, Burkman RT, Weiss LK, Spirtas R (2005) Reproductive factors and subtypes of breast cancer defined by hormone receptor and histology. *Br J Cancer* **93**(3): 364–371.

Vatten LJ, Kvinnsland S (1990) Cigarette smoking and risk of breast cancer: a prospective study of 24,329 Norwegian women. *Eur J Cancer* **26**(7): 830–833.

Xue F, Willett WC, Rosner BA, Hankinson SE, Michels KB (2011) Cigarette smoking and the incidence of breast cancer. *Arch Intern Med* **171**(2): 125–133.

Young Jr JL, Roffers SD, Ries LAG, Fritz AG, Hurlbut AA (eds) (2001) *SEER Summary Staging Manual - 2000: Codes and Coding Instructions*. National Cancer Institute: Bethesda, MD, NIH Pub. No. 01-4969).

Supplementary Information accompanies this paper on British Journal of Cancer website (http://www.nature.com/bjc)