

**Keywords:** breast cancer incidence; ethnicity; breast cancer risk factors

# Ethnic differences in breast cancer incidence in England are due to differences in known risk factors for the disease: prospective study

T Gathani<sup>\*1,2</sup>, R Ali<sup>1</sup>, A Balkwill<sup>1</sup>, J Green<sup>1</sup>, G Reeves<sup>1</sup>, V Beral<sup>1</sup>, KA Moser<sup>1</sup> and on behalf of the Million Women Study Collaborators<sup>3</sup>

<sup>1</sup>Cancer Epidemiology Unit, University of Oxford, Richard Doll Building, Roosevelt Drive, Oxford OX3 7LF, UK and <sup>2</sup>Department of Breast Surgery, Oxford University Hospitals NHS Trust, Oxford, UK

**Background:** In the United Kingdom, breast cancer incidence is lower in South Asian and Black women than in White women, but the extent to which this is due to known risk factors is unknown. In a large prospective study, we describe breast cancer incidence by ethnicity, before and after adjustment for known risk factors for the disease.

**Methods:** Women were recruited into the Million Women Study in 1996–2001, when information on reproductive and lifestyle factors known to influence the risk of breast cancer was obtained. Ethnicity was determined from study questionnaires and hospital admission data. Cox regression models were used to calculate adjusted relative risks (RR) for incident breast cancer in South Asians and Blacks compared with Whites.

**Results:** Analyses included 5877 South Asian, 4919 Black, and 1 038 144 White women in England. The prevalence of 8 out of the 9 risk factors for breast cancer examined, differed substantially by ethnicity ( $P < 0.001$  for each), such that South Asian and Black women were at a lower risk of the disease than White women. During 12.2 years of follow-up incident breast cancer occurred in 217 South Asians, 180 Blacks, and 45 191 Whites. As expected, breast cancer incidence was lower in South Asians (RR = 0.82, 95% CI 0.72–0.94) and Blacks (RR = 0.85, 0.73–0.98) than in Whites when the analyses were adjusted only for age and region of residence. However, after additional adjustment for the known risk factors for the disease, breast cancer incidence was similar to that of Whites, both in South Asians (0.95, 0.83–1.09) and in Blacks (0.91, 0.78–1.05).

**Conclusion:** South Asian and Black women in England have lower incidence rates of breast cancer than White women, but this is largely, if not wholly, because of differences in known risk factors for the disease.

Breast cancer is the most common cancer in women in England, with over 41 000 new cases diagnosed annually (Office for National Statistics, 2012a). People of South Asian ethnicity represent the largest ethnic minority in England and Wales (5% of total population) and Blacks constitute the second largest ethnic minority group (3%) (Office for National Statistics, 2012b).

South Asian and Black women are known to have a lower incidence of breast cancer than White women, both in the United Kingdom (Farooq and Coleman, 2005; Jack *et al*, 2009; National

Cancer Intelligence Network, 2009; Ali *et al*, 2010; Downing *et al*, 2011) and in the United States of America (Goggins and Wong, 2009; DeSantis *et al*, 2011; Siegel *et al*, 2012). The reasons for the lower incidence of breast cancer in South Asian and Black women are largely unexplained, as studies with sufficiently detailed information to examine these differences are lacking.

Breast cancer incidence is known to be affected by reproductive, lifestyle, and other factors (Key *et al*, 2001). Factors that are known to have independent effects reducing the subsequent risk of breast

\*Correspondence: Dr T Gathani; E-mail: toral.gathani@ceu.ox.ac.uk

<sup>3</sup>See Appendix.

cancer include an older age at menarche (Collaborative Group on Hormonal Factors in Breast Cancer, 2012), increased parity and duration of breastfeeding (Collaborative Group on Hormonal Factors in Breast Cancer, 2002b), shorter stature (Green *et al*, 2011), and lower body mass index (Reeves *et al*, 2007). Factors known to have independent effects increasing the subsequent risk of breast cancer include greater alcohol consumption (Collaborative Group on Hormonal Factors in Breast Cancer, 2002a), use of menopausal hormone therapy (The Million Women Study Collaborators, 2003), and a family history of the disease (Collaborative Group on Hormonal Factors in Breast Cancer, 2001).

Little is known, however, about how the prevalence of these factors among different ethnic groups in the United Kingdom affects comparisons of breast cancer incidence by ethnicity. Using data from a large prospective study, we report here on the prevalence of risk factors for breast cancer among South Asian, Black, and White women in England and on the incidence of breast cancer in these groups, before and after adjusting for the risk factors.

## MATERIALS AND METHODS

**Study population.** The Million Women Study is a prospective cohort study of 1.3 million women who were recruited when they were invited for screening by the National Health Service (NHS) Breast Screening Programme in England and Scotland from 1996 to 2001. At this time, routine invitations to attend breast screening were provided to all women registered with a general practitioner and aged between 50 and 64 years. Over 99% of women in the study cohort have undergone mammographic screening. At recruitment, women completed a questionnaire asking about socio-demographic factors, their height, weight, alcohol consumption, use of hormonal therapies, reproductive and medical history, and place of birth. Full details of the study design and methods are described elsewhere (The Million Women Study Collaborators, 2003) and questionnaires can be viewed at [www.millionwomenstudy.org](http://www.millionwomenstudy.org). Study participants have been flagged on the NHS Central Register, so that study investigators are routinely notified of cancer registrations and deaths, including the date of each event and information on the cancer site or cause of death, coded according to the International Classification of Diseases, 10th revision (World Health Organisation, 1992). To date, we have

complete follow-up on ~99% of the study population. All participants gave written consent for inclusion in the study, and approval was obtained from the Oxford and Anglia Multi-Centre Research Ethics Committee.

**Determination of ethnicity.** We used two sources of information to ascertain ethnicity. All Million Women Study participants were asked at recruitment about their place of birth and this was used to derive country of birth. At resurvey 3 years later they were asked 'To which ethnic group do you consider you belong?'; most women were given a choice of 'White', 'Asian', 'Black – Caribbean, African etc.' and 'Other' but early versions of the study questionnaire included the option 'South Asian' instead of 'Asian'.

Information on ethnicity was also obtained from hospital admission records (Hospital Episode Statistics, HES) to which study participants recruited in England were linked, using their NHS number and date of birth. Hospital data in England have collected information on ethnicity since 1995 (Department of Health and NHS Health and Social Care Information Centre, 2005). In Scotland, information about ethnicity is not routinely collected in hospital data, and so it was not possible to assess the consistency of the ethnicity information provided by women recruited in Scotland ( $n = 117\,120$ ). These women were not included in the analyses.

In assigning ethnicity, the ethnicity reported in the Million Women Study took priority over that reported in HES data which was only used to supplement the Million Women Study information. Table 1 shows in detail how ethnic groups were defined in the analysis. To assess the robustness of these definitions, we examined the consistency of ethnicity information for the women who had ethnicity recorded in more than one data source.

**Analysis.** We compared characteristics of women by their ethnicity, for South Asian, Black, and White women. Women with unknown ethnicity ( $n = 152\,323$ ) were omitted from the analysis as were women with a cancer (other than non-melanoma skin cancer, ICD-10, C44) registration before recruitment ( $n = 37\,808$ ). Cox regression models, with age as the underlying variable, were used to calculate relative risks (RRs) of breast cancer (ICD-10, C50) in each of South Asian and Black women compared with White women. Eligible women contributed person-years from the date of recruitment until the date of first registration of any cancer (other than non-melanoma skin cancer), emigration, death

Table 1. Definitions of ethnicity used in the analysis

Ethnicity	Criteria (fulfil any of the below)
South Asian	(a) Reported on the Million Women Study questionnaire that ethnic group was 'South Asian', or (b) reported on the Million Women Study questionnaire that ethnic group was 'Asian' and that they were not born in China, Hong Kong, Malaysia, Japan, the Philippines, Thailand, Burma, Singapore, Vietnam or Iran, or (c) ethnic group could not be determined from the Million Women Study ('Other', not stated or missing) but recorded as being of Indian, Pakistani, or Bangladeshi ethnicity on one or more HES records
Black	(a) Reported on the Million Women Study questionnaire that ethnic group was 'Black', or (b) ethnic group could not be determined from the Million Women Study but recorded as being of 'Black' ethnicity on one or more HES records
White	(a) Reported on the Million Women Study questionnaire that ethnic group was 'White', or (b) ethnic group could not be determined from the Million Women Study but recorded as White on all their HES records with stated ethnicity
Other <sup>a</sup>	(a) 'Other', not stated or missing ethnicity according to the Million Women Study and who could not be assigned to South Asian, Black, or White ethnic groups from HES records, or (b) reported on the Million Women Study questionnaire that ethnic group was 'Asian' and that they were born in China, Hong Kong, Malaysia, Japan, the Philippines, Thailand, Burma, Singapore, Vietnam, or Iran (i.e., not South Asian), or (c) inconsistent recording of ethnicity

Abbreviation: HES = Hospital Episode Statistics.

<sup>a</sup>The largest specific ethnic group included under 'Other' was Chinese women, defined as: (a) reported on the Million Women Study questionnaire that ethnic group was 'Asian' and that they were born in China, Hong Kong, Malaysia or Singapore, or (b) ethnic group could not be determined from the Million Women Study but recorded as being of 'Chinese' ethnicity on one or more HES records.

from any cause or the end of follow-up, whichever was the earliest. The end of follow-up was 31 December 2010 for cancer registry regions East Anglia, Thames, Trent, North West (Mersey), and North West (Manchester/Lancashire); 30 June 2011 for Oxford and North Yorkshire regions, and 31 December 2011 for South West and West Midlands regions.

Initially, analyses were adjusted for age and region of residence (9 regions, representing the areas covered by cancer registries in England). We then investigated the effects of additional adjustments for nine risk factors for breast cancer, one at a time, and finally for all nine factors simultaneously. These risk factors were an index of socioeconomic status (in quintiles, of the Townsend score for area of residence; Townsend *et al*, 1988), age at menarche (<13, 13–14, and 15+), parity and age at first birth (nulliparous, age<25 at first birth and parity 1–2, age at first birth 25+ and parity 1–2, age at first birth <25 and parity ≥3, age at first birth 25+ and parity ≥3) and breastfeeding (ever, never), body mass index (<25, 25 to <30, 30+ kg m<sup>-2</sup>), height (<160 cm, 160 to <165 cm, 165+ cm), alcohol consumption (non-drinker, <7 drinks per week, 7+ drinks per week), use of menopausal hormone therapy (current, never, past), and family history of breast cancer (having or not having an affected first-degree relative). Where there were missing values for a woman in any of the adjustment variables, they were assigned to a separate category for that adjustment. Fewer than 2% of women had missing values for any single adjustment variable, and they were assigned to a separate category for that adjustment, so that exactly the same women were included in each analysis.

**RESULTS**

Using both Million Women Study and HES ethnicity data, we were able to assign ethnicity to 87% (1 056 992/1 209 315) of the women recruited in England who did not have a prior cancer. Of these 1 056 992 women, 5877 were South Asian, 4919 were Black, and 1 038 144 were White (Table 2). Their average ages at recruitment were 55.3 (standard deviation (s.d.) 4.6), 56.6 (s.d. 4.7), and 56.2 (s.d. 4.9) years, respectively.

Table 2 shows the socio-demographic and lifestyle characteristics of women at recruitment. Large differences by ethnic group were seen for all the factors examined (*P*<0.001 for each comparison). For example, almost two-thirds of the South Asian and Black women, compared with one-third of the White women, were in the lowest deprivation tertile. South Asian and Black women had more children than White women, were more likely to breastfeed them, and were less likely to drink alcohol and to have a first-degree relative with breast cancer. South Asian and Black women were less likely to use hormone therapy for the menopause compared with White women, but the overall duration of use among current users was similar across the three ethnic groups (on average about 5 years). With the exception of body mass index, the ethnic differences in characteristics shown in Table 2 would be expected to reduce breast cancer risk in South Asian and Black women compared with White women.

During an average follow-up period of 12.2 years, there were 217 incident breast cancers in South Asians, 180 in Blacks, and 45 191 in Whites. The separate effect of each of the variables on the RRs for breast cancer in South Asian and Black women compared with White women is shown in Table 3. For South Asian compared with White women, the RR of breast cancer was significantly reduced when the risk estimates were adjusted by age and region only (RR = 0.82, 95% CI 0.72–0.94; *P* = 0.004). Additional adjustment, separately by each of the nine above-mentioned risk factors brought some of the RR estimates slightly closer to 1.0, in particular those for childbearing and breastfeeding, height, alcohol consumption, and use of menopausal hormone therapy. After simultaneous adjustment for all factors, the risk of breast cancer in South Asian women was similar to that of White women (RR = 0.95, 95% CI 0.83–1.09; *P* = 0.5).

For Black compared with White women, the RR of breast cancer was also significantly reduced when the risk estimates were adjusted by age and region only (RR = 0.85, 95%CI 0.73–0.98; *P* = 0.03). Additional adjustment separately by the nine above-mentioned risk factors brought some RR estimates slightly closer to 1.0 (except for adjustment for body mass index, as Black women have a higher average body mass index than White women, Table 2). After simultaneous adjustment for all factors, the risk of

**Table 2.** Characteristics at recruitment of South Asian, Black, and White women in England, and details of follow-up

	Ethnic group		
	South Asian	Black	White
Number of women	5877	4919	1 038 144
<b>Characteristics at recruitment*</b>			
% (n) in most deprived tertile	64 (3765)	69 (3385)	33 (338 392)
Age at menarche (years), mean (s.d.)	13.7 (1.6)	13.6 (1.8)	13.0 (1.6)
Number of children (all women), mean (s.d.)	2.7 (1.6)	2.9 (1.8)	2.1 (1.2)
Age (years) at first birth (parous women), mean (s.d.)	23.8 (4.9)	22.2 (4.7)	23.8 (4.3)
% (n) ever breastfed a child (parous women)	85 (3050)	83 (2486)	69 (506 547)
Body mass index (kg m <sup>-2</sup> ), mean (s.d.)	26.3 (4.6)	28.1 (5.2)	26.2 (4.7)
Height (cm), mean (s.d.)	157.7 (6.7)	161.8 (7.1)	162.1 (6.7)
% (n) non-drinker of alcohol	75 (4266)	38 (1829)	23 (240 220)
% (n) current user of menopausal hormone therapy	22 (1261)	29 (1350)	35 (355 731)
% (n) with first-degree relative with breast cancer	5 (312)	7 (340)	10 (98 709)
<b>Follow-up</b>			
Women years of follow-up (1000 s)	69.9	56.3	12 127.4
Number of incident breast cancers	217	180	45 191

\**P*<0.001 for each comparison.

**Table 3.** Relative risk (RR) for breast cancer in South Asian and Black compared with White women in England, adjusted by various factors

	South Asian vs White RR (95% CI)	Black vs White RR (95% CI)
Number of incident breast cancers	South Asian women: 217	Black women: 180
<b>Adjusted by age and region only</b>	<b>0.82 (0.72–0.94)</b>	<b>0.85 (0.73–0.98)</b>
<b>Additional adjustment, separately by:</b>		
Deprivation	0.83 (0.72–0.95)	0.86 (0.74–0.99)
Age at menarche	0.83 (0.72–0.94)	0.85 (0.74–0.99)
Childbearing and breastfeeding history	0.84 (0.74–0.96)	0.88 (0.76–1.02)
Body mass index	0.82 (0.72–0.94)	0.84 (0.72–0.97)
Height	0.86 (0.75–0.98)	0.85 (0.74–0.99)
Alcohol consumption	0.86 (0.75–0.99)	0.87 (0.75–1.01)
Use of menopausal hormone therapy	0.85 (0.75–0.98)	0.87 (0.75–1.00)
First-degree relative with breast cancer	0.84 (0.73–0.96)	0.86 (0.74–0.99)
<b>Simultaneous adjustment by all the above</b>	<b>0.95 (0.83–1.09)</b>	<b>0.91 (0.78–1.05)</b>

breast cancer in Black women was similar to that of White women (RR = 0.91, 95% CI 0.78–1.05;  $P = 0.2$ ).

Among the 1 056 992 women included in the analyses, 8052 women were classified as being of 'Other' ethnicity, and among these women 345 incident breast cancers were recorded. The largest specific ethnic group included in this 'Other' category was Chinese women (976) among whom 30 incident breast cancers were recorded. The associated RRs for breast cancer were 0.70 (0.49–1.01) adjusted for age and region only, and 0.75 (0.53–1.08) after simultaneous adjustment for all the factors listed in Table 3.

The HES data distinguish between South Asian and Chinese ethnicities, but the Million Women Study does not. Therefore, it was possible to validate the use of country of birth to assign women to South Asian on the basis of self-reported ethnicity and country of birth where no HES record was available. Of the women who were recorded in the Million Women Study as Asian and born in Pakistan, Bangladesh, or India, 89% were recorded as having a South Asian ethnicity in HES records. By contrast, of the women recorded in the Million Women Study as Asian and born in China, Japan, Hong Kong, Malaysia, the Philippines, Thailand, Burma, Singapore, or Iran, only 8% were recorded as having a South Asian ethnicity in HES records but 80% were recorded as Chinese. Among women who according to Million Women Study records were Black and born in the Caribbean, United States, or Africa, 89% were recorded as Black in HES records.

## DISCUSSION

These findings indicate that the lower incidence rates of breast cancer seen in South Asian and Black as compared with White women in England are largely, if not wholly, because of differences in known risk factors for the disease. Once adjustments were made for risk factors including age at menarche, height, childbearing and breastfeeding history, alcohol consumption, and use of menopausal hormone therapy, South Asian and Black women were shown to have similar breast cancer risks to White women.

To our knowledge, this is the first study in the United Kingdom able to compare breast cancer incidence in South Asian, Black, and White women, adjusting for most known risk factors for the

disease. We found large differences between ethnic groups in the known risk factors; such differences have been reported previously in regional studies in England (Modood *et al*, 1997; dos Santos Silva *et al*, 2003; Velikova *et al*, 2004; Farooq and Coleman, 2005; McCormack *et al*, 2008; Jack *et al*, 2009). When our findings on breast cancer incidence were adjusted only by women's age and region of residence, South Asians and Blacks were found to have lower risks of breast cancer than Whites, which is consistent with results from cancer registries, where adjustments could only be made by age and geographical location, and sometimes by socioeconomic status (Farooq and Coleman, 2005; National Cancer Intelligence Network, 2009; Ali *et al*, 2010).

Examining the differences in breast cancer incidence in women of different ethnic groups is of interest. A study in the United States (US) found that the lower breast cancer incidence observed in Hispanics, American Indian/Alaskan Native, and Asian/Pacific Islanders, but not African Americans, compared with White women were accounted for by known risk factors (Chlebowski *et al*, 2005). However, the ethnic composition of the US population bears little resemblance to that of the United Kingdom. Both have their unique historical migration patterns, and as a consequence of this, findings for the US, or indeed elsewhere, have limited relevance to the UK setting. Most women from minority ethnic groups in the Million Women Study are first-generation migrants from South Asia, the Caribbean, and Africa (Office for National Statistics, 2010).

An important strength of this work is its use of data from a prospective study of over one million women. This means there are sufficient numbers of women in the age range and ethnic groups considered to give a robust analysis. As, in each of the minority ethnic groups considered here, only a small proportion are aged 50 and above, a large data set is required to conduct a study of ethnic differences in this age range. And even in this large data set the South Asian and Black groups cannot be further differentiated as the number of incident breast cancers in the component groups would be insufficient to give a robust analysis. Thus, any differences within the South Asian and Black ethnic groups could not be examined by this study. Other strengths of this study are that information on risk factors for breast cancer is collected before cancer diagnosis, and women are followed up for cancer incidence and death using the comprehensive national registration system.

The limitations of the study include a potential recruitment bias in that recruitment into the Million Women Study covered about 50% of areas covered by the NHS Breast Screening Programme in England. As London, with its high proportion of the population from ethnic minorities, is underrepresented in the Million Women Study, the prevalence of ethnic minorities is lower in this study than in the overall population of England. A representative national survey in Great Britain found no significant differences by ethnicity in the proportion who attended for screening mammography (Moser *et al*, 2009), although in some localities the uptake of mammographic screening has been reported to be lower for South Asian and Black women than for White women (Szczerpura *et al*, 2008; Renshaw *et al*, 2010). Other limitations include a lack of genetic information on participants in the Million Women Study. The Million Women Study is restricted to women aged 50–64 at recruitment to the study; the average follow-up of around 12 years results in women aged 50–75 being included in the analysis. This is a study of breast cancer in older women, and comparisons with other age groups are not possible.

In the United Kingdom, ethnicity is taken as self-defined (Office for National Statistics, 2003). A further strength of this work is that we used self-reported data on ethnicity from two independent sources, the Million Women Study and hospital admission (HES) data. Million Women Study data were taken as the primary source of information and, where necessary, were supplemented by

ethnicity as recorded in HES records. Specifically, HES ethnicity data were drawn on for about one-third of the study population who did not respond at 3-year resurvey (when the questionnaire asked about ethnic group) or whose response at resurvey could not be classified to an ethnic group (2% of responders). The completeness of recording of ethnicity in HES has improved recently; the proportion of episodes where the ethnic code was recorded as 'Not known' or 'Not stated' declined from 24% in 2004–2005 to 9% in 2009–2010 (HES Online, 2011). Our findings also suggest that the HES coding of ethnicity is fairly reliable, in that there was good agreement for those defined as South Asian and Black where ethnicity information was available from both data sets.

The results of this study suggest that the ethnic differences in breast cancer incidence currently observed in England are largely due to ethnic differences in reproductive and lifestyle factors. Previous meta-analyses reported no significant variation by women's ethnicity of the effects of equivalent exposures to breastfeeding and childbearing (Collaborative Group on Hormonal Factors in Breast Cancer, 2002b), alcohol consumption (Collaborative Group on Hormonal Factors in Breast Cancer, 2002a), use of menopausal hormone therapy (Collaborative Group on Hormonal Factors in Breast Cancer, 1997), and of a family history of breast cancer (Collaborative Group on Hormonal Factors in Breast Cancer, 2001) on breast cancer risk. Acculturation among migrants to the United Kingdom is known to be associated with changes in health behaviours such as alcohol consumption and breastfeeding (Hawkins *et al*, 2008), and reproductive factors, including parity and age at first birth, differ between first- and second-generation migrants (Leon and Moser, 2012). Consequently, likely changes in exposures to risk factors for breast cancer by younger generations may well alter ethnic differences in breast cancer incidence in the future.

## ACKNOWLEDGEMENTS

We thank the women who participated in the Million Women Study. This work was supported by Cancer Research UK (C570/A11692) and UK Medical Research Council (G0700474). The funders did not influence the conduct of the study, the preparation of this report, or the decision to publish. The authors had full access to all the data in the study and had final responsibility for the decision to submit for publication. The study is registered with the NHS National Institute of Health Research Portfolio (study number 6862). The corresponding author (Toral Gathani) has the right to grant on behalf of all authors and does grant on behalf of all authors, a worldwide licence to the Publishers and its licensees in perpetuity, in all forms, formats and media (whether known now or created in the future), to (i) publish, reproduce, distribute, display, and store the Contribution, (ii) translate the Contribution into other languages, create adaptations, reprints, include within collections and create summaries, extracts and/or, abstracts of the Contribution, (iii) create any other derivative work(s) based on the Contribution, (iv) to exploit all subsidiary rights in the Contribution, (v) the inclusion of electronic links from the Contribution to third party material wherever it may be located; and, (vi) licence any third party to do any or all of the above.

## CONFLICT OF INTEREST

The authors declare no conflict of interest.

## AUTHOR CONTRIBUTIONS

All authors participated in the design and conduct of the study and read and approved the final manuscript. TG and AB are the guarantors.

## ETHICAL APPROVAL

The Million Women Study has been approved by the Oxford and Anglia (now East of England-Cambridge South) Multi-centre Research Ethics Committee (Ref: REC 97/5/01), and all study participants gave signed consent to be included. Access and linkage to hospital records was approved by the Information Centre for Health and Social Care in England.

## DATA SHARING

All information provided is stored in accordance with the Data Protection Act (Office of Data Protection Registrar registration No K3039784). Only the study team has access to computerised data, *via* passwords (see the Million Women Study protocol at [www.millionwomenstudy.org](http://www.millionwomenstudy.org)).

## REFERENCES

- Ali R, Barnes I, Kan SW, Beral V (2010) Cancer Incidence in British Indians and British whites in Leicester, 2001–2006. *Br J Cancer* **103**: 143–148.
- Chlebowski RT, Chen Z, Anderson GL, Rohan T, Aragaki A, Lane D, Dolan NC, Paskett ED, McTiernan A, Hubbell FA, Adams-Campbell LL, Prentice R (2005) Ethnicity and breast cancer: factors influencing differences in incidence and outcome. *J Natl Cancer Inst* **97**(6): 439–448.
- Collaborative Group on Hormonal Factors in Breast Cancer (1997) Breast cancer and hormone replacement therapy: collaborative reanalysis of data from 51 epidemiological studies of 52,705 women with breast cancer and 108,411 women without breast cancer. *Lancet* **350**(9084): 1047–1059.
- Collaborative Group on Hormonal Factors in Breast Cancer (2001) Familial breast cancer: collaborative reanalysis of individual data from 52 epidemiological studies including 58,209 women with breast cancer and 101,986 women without the disease. *Lancet* **358**(9291): 1389–1399.
- Collaborative Group on Hormonal Factors in Breast Cancer (2002a) 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.
- Collaborative Group on Hormonal Factors in Breast Cancer (2002b) Breast cancer and breastfeeding: collaborative reanalysis of individual data from 47 epidemiological studies in 30 countries, including 50302 women with breast cancer and 96 973 women without the disease. *Lancet* **360**(9328): 187–195.
- Collaborative Group on Hormonal Factors in Breast Cancer (2012) Menarche, menopause and breast cancer risk: individual participant meta-analysis including 118 964 women with breast cancer from 117 epidemiological studies. *Lancet Oncol* **13**: 1141–1151.
- Department of Health and NHS Health and Social Care Information Centre (2005) Practical guide to ethnic monitoring in the NHS and social care [http://www.dh.gov.uk/en/Publicationsandstatistics/Publications/PublicationsPolicyAndGuidance/DH\\_4116839](http://www.dh.gov.uk/en/Publicationsandstatistics/Publications/PublicationsPolicyAndGuidance/DH_4116839).
- DeSantis C, Siegel R, Bandi P, Jemal A (2011) Breast Cancer Statistics, 2011. *CA Cancer J Clin* **61**: 409–418.
- dos Santos Silva I, Mangtani P, De Stavola BL, Bell J, Quinn M, Mayer D (2003) Survival from breast cancer among South Asian and non-South Asian women resident in South East England. *Br J Cancer* **89**: 508–512.
- Downing A, West RM, Gilthorpe MS, Lawrence G, Forman D (2011) Using routinely collected health data to investigate the association between ethnicity and breast cancer incidence and survival: what is the impact of missing data and multiple ethnicities? *Ethn Health* **16**(3): 201–212.
- Farooq S, Coleman MP (2005) Breast cancer survival in South Asian women in England and Wales. *J Epidemiol Commun Health* **59**: 402–406.

- Goggins WB, Wong G (2009) Cancer among Asian Indians/Pakistanis living in the United States: low incidence and generally above average survival. *Cancer Causes Control* **20**: 635–643.
- Green J, Cairns BJ, Cassabonne D, Wright FL, Reeves G, Beral V (2011) Height and cancer incidence in the Million Women Study: prospective cohort, and meta-analysis of prospective studies for height and total cancer risk. *Lancet Oncol* **12**(8): 785–794.
- Hawkins SS, Lamb K, Cole TJ, Law C. and Millennium Cohort Study Child Health Group (2008) Influence of moving to the UK on maternal health behaviours: prospective cohort study. *Br Med J* **336**(7652): 1052–1055.
- HES Online (2011) *How Good Is HES Ethnic Coding and Where do the Problems lie?* Health and Social Care Information Centre: Leeds.
- Jack RH, Davies EA, Moller H (2009) Breast cancer incidence, stage, treatment and survival in ethnic groups in South East England. *Br J Cancer* **100**: 545–550.
- Key T, Verkasalo P, Banks E (2001) Epidemiology of breast cancer. *Lancet Oncol* **2**: 133–140.
- Leon DA, Moser K (2012) Low birth weight persists in South Asian babies born in England and Wales regardless of maternal country of birth. Slow pace of acculturation, physiological constraint or both? Analysis of routine data. *J Epidemiol Commun Health* **66**(6): 544–551.
- McCormack V, Perry N, Vinnicombe S, dos Santos Silva I (2008) Ethnic variations in mammographic density: a British multiethnic longitudinal study. *Am J Epidemiol* **168**: 412–421.
- Modood T, Berthoud R, Lakey J, Nazroo J, Smith P, Virdee S, Beishon S (1997) *Ethnic Minorities in Britain: Diversity and Disadvantage*. Policy Studies Institute: London, England.
- Moser K, Patnick J, Beral V (2009) Inequalities in reported use of breast and cervical screening in Great Britain: analysis of cross sectional survey data. *Br Med J* **338**: 2025–2034.
- National Cancer Intelligence Network (2009) 'Cancer Incidence and Survival by Major Ethnic Group, England, 2002–2006 (online)'. <http://www.ncin.org.uk> (accessed June 2013).
- Office for National Statistics (2003) *Ethnic Group Statistics: A Guide for the Collection and Classification of Ethnicity Data*. ONS: Newport.
- Office for National Statistics (2010) 2001 Census of England and Wales: Country of Birth by Ethnic Groups by Sex and Age Group (Previously Commissioned Table C1151) <http://www.ons.gov.uk/ons/guide-method/census/census-2001/data-and-products/data-and-product-catalogue/commissioned-output/commissioned-tables/index.html>.
- Office for National Statistics (2012a) Breast Cancer: Incidence, Mortality and Survival, 2010. <http://www.ons.gov.uk/ons/rel/cancer-unit/breast-cancer-in-england/2010/sum-1.html> (accessed June 2013).
- Office for National Statistics (2012b) 2011 Census: Key Statistics for Local Authorities in England and Wales Table KS201EW: Ethnic Group, Local Authorities in England and Wales. <http://www.ons.gov.uk> (accessed June 2013).
- Reeves G, Pirie K, Beral V, Green J, Spencer E, Bull D (2007) Cancer incidence and mortality in relation to body mass index in the Million Women Study: cohort study. *Br Med J* **335**(7630): 1134–1139.
- Renshaw C, Jack RH, Dixon S, Moller H, Davies EA (2010) Estimating attendance for breast cancer screening in ethnic groups in London. *BMC Public Health* **10**: 157–165.
- Siegel R, Naishadham D, Jemal A (2012) Cancer Statistics, 2012. *CA Cancer J Clin* **62**(1): 10–29.
- Szczepura A, Price C, Gumber A (2008) Breast and bowel cancer screening uptake patterns over 15 years for UK South Asian ethnic minority populations, corrected for differences in socio-demographic characteristics. *BMC Public Health* **8**: 346–361.
- The Million Women Study Collaborators (2003) Breast cancer and hormone replacement therapy in the Million Women Study. *Lancet* **362**: 419–427.
- Townsend P, Phillimore P, Beattie A (1988) *Health and Deprivation: Inequality and the North*. Croom Helm: London.
- Velikova G, Booth L, Johnston C, Forman D, Selby P (2004) Breast cancer outcomes in South Asian population of West Yorkshire. *Br J Cancer* **90**: 1926–1932.
- World Health Organisation (1992) International Statistical Classification of Diseases and Related Health Problems, 10th revision. World Health Organisation: Geneva.

This work is published under the standard license to publish agreement. After 12 months the work will become freely available and the license terms will switch to a Creative Commons Attribution-NonCommercial-Share Alike 3.0 Unported License.

## APPENDIX

### Million Women Study Collaborators

#### The Steering Committee are:

Emily Banks, Valerie Beral, Ruth English, Jane Green, Julietta Patnick, Richard Peto, Gillian Reeves, Martin Vessey, and Matthew Wallis.

#### The Million Women Study Co-ordinating Centre staff are as follows:

Hayley Abbiss, Simon Abbott, Naomi Allen, Miranda Armstrong, Angela Balkwill, Emily Banks, Vicky Benson, Valerie Beral, Judith Black, Anna Brown, Benjamin Cairns, Karen Canfell, Dexter Canoy, James Chivenga, Barbara Crossley, Francesca Crowe, Dave Ewart, Sarah Ewart, Lee Fletcher, Toral Gathani, Laura Gerrard, Adrian Goodill, Jane Green, Lynden Guiver, Isobel Lingard, Sau Wan Kan, Oksana Kirichek, Mary Kroll, Nicky Langston, Bette Liu, Maria-Jose Luque, Kath Moser, Lynn Pank, Kirstin Pirie, Gillian Reeves, Keith Shaw, Emma Sherman, Evie Sherry-Starmer, Helena Strange,

Sian Sweetland, Alison Timadger, Sarah Tipper, Ruth Travis, Lucy Wright, Owen Yang, and Heather Young.

#### The following NHS Breast Screening Centres took part in the recruitment and breast screening follow-up for the Million Women Study:

Avon, Aylesbury, Barnsley, Basingstoke, Bedfordshire and Hertfordshire, Cambridge and Huntingdon, Chelmsford and Colchester, Chester, Cornwall, Crewe, Cumbria, Doncaster, Dorset, East Berkshire, East Cheshire, East Devon, East of Scotland, East Suffolk, East Sussex, Gateshead, Gloucestershire, Great Yarmouth, Hereford and Worcester, Kent, Kings Lynn, Leicestershire, Liverpool, Manchester, Milton Keynes, Newcastle, North Birmingham, North East Scotland, North Lancashire, North Middlesex, North Nottingham, North of Scotland, North Tees, North Yorkshire, Nottingham, Oxford, Portsmouth, Rotherham, Sheffield, Shropshire, Somerset, South Birmingham, South East Scotland, South East Staffordshire, South Derbyshire, South Essex, South Lancashire, South West Scotland, Surrey, Warrington Halton St Helens and Knowsley, Warwickshire Solihull and Coventry, West Berkshire, West Devon, West London, West Suffolk, West Sussex, Wiltshire, Winchester, Wirral, and Wycombe.