THE MILLION WOMEN STUDY: UPDATED PROTOCOL

Study Title:	The Million Women Study
Short title:	The Million Women Study
Study website:	www.millionwomenstudy.org
Ethics Ref:	REC 97/5/001 [original study protocol approved MREC May 1997]
Date and Version No:	May 2018 V 2.3
Chief Investigators:	Professor Dame Valerie Beral Professor Jane Green Professor Gillian Reeves Cancer Epidemiology Unit Nuffield Department of Population Health University of Oxford, UK jane.green@ceu.ox.ac.uk
Sponsor:	University of Oxford
Funders:	Cancer Research UK Medical Research Council
Data Controller	University of Oxford
Data Processor	Cancer Epidemiology Unit, Nuffield Department of Population Health, University of Oxford

No conflicts of interest

Document history

Version number	**Date**		
1.0	May 1997	Original study protocol Approved MREC 97/5/001	Valerie Beral
1.1 1.2 1.3 1.4	April 2015 April 2016 May 2016 October 2016	Approved additions to protocol for amendments 27,29,30 and 31	Jane Green Lynden Guiver
2.0	February 2017	Updated study protocol	Jane Green Gillian Reeves Valerie Beral
2.1	August 2017	Revisions to protocol for amendment 33 New version1.4 of MWS data sharing policy Table 2 revised	Jane Green Lynden Guiver
2.2	November 2017	Revisions to protocol for Amendment 34	Jane Green Lynden Guiver
2.3	May 2018	Revision to Data Access Policy	Jane Green

TABLE OF CONTENTS **PAGE NO**

1. ABBREVIATIONS

CEU	Cancer Epidemiology Unit
CPRD	Clinical Practice Research Datalink
HT, HRT	Hormone therapy for menopause, Hormone Replacement Therapy
HSCIC	Health and Social Care Information Centre
ISD Scotland	Information Services Division Scotland
NDPH	Nuffield Department of Population Health
NHS	National Health Service
ONS	Office for National Statistics
(M)REC	(Multi-centre) Research Ethics Committee

2. SYNOPSIS

Study Title	The Million Women Study	
Internal ref. no. / short title	The Million Women Study	
Study Design	Observational prospective cohort study	
Study Participants	1.3 million UK women aged 50-64 years at recruitment	
Planned Study Period	Ongoing; participants recruited in 1996-2001	
	Objectives	**Outcome Measures**
Primary	To provide reliable information on the effects of potentially modifiable factors that affect women's health in middle age and in old age. To investigate risks and burden of fatal and non-fatal cancer and vascular, neurodegenerative, musculoskeletal and other diseases in relation to demographic, social, health and behavioural factors.	Incidence of cancer and site of primary tumour Incidence of death and cause of mortality Incidence of and rates of hospital admissions for vascular, neurodegenerative, musculoskeletal and other diseases

3. SUMMARY

The Million Women Study is an open-ended prospective study of the health of women in England and Scotland, and includes 1.3 million women recruited between 1996 and 2001 at age 50-64 years through National Health Service (NHS) Breast Screening Centres. The aim of the study is to provide

reliable information on the effects of potentially modifiable factors that affect women's health in middle age and in old age.

This protocol updates the original protocol approved at the start of the study some 20 years ago (REC 97/5/001, Appendix 1), and describes study development through 31 approved amendments to the original protocol (Appendix 2). A separate protocol was approved in 2004 for collection of biological samples (Appendix 3).

The initial stimulus was to obtain robust prospective information on the risk of breast cancer associated with use of different types of menopausal hormone therapy (HT), and we have shown that use of oestrogen-progestogen preparations causes much greater increases in the risk of breast cancer than oestrogen-only preparations; further analyses have described the overall impact of HT use on risk for breast, ovarian and endometrial cancers. When planning the necessary large-scale prospective study, an equally important aim was to obtain reliable information on the effects of other potentially modifiable factors affecting women's health, and we have published the results of investigations of risk for vascular, musculoskeletal and neurodegenerative disease, as well as for various cancers, in relation to smoking, obesity, alcohol consumption, and body size.

The cohort has continuing follow-up through postal and online re-surveys, to update exposure measures, and through electronic linkage to routinely-collected NHS records for disease outcomes, including death, cancer registration and hospital admissions databases. After almost 20 years follow up only 1.5% of women in the cohort have been lost to follow-up, 13% have died, 15% had a cancer registration and 85% have had at least one hospital admission. Blood samples for genetic and biochemical analyses have been collected for some 60,000 women. Through continuing follow-up and collaboration with other researchers we plan to maximise the value to public health of this rich data resource for many years to come.

4. BACKGROUND AND RATIONALE

This protocol updates that written at the start of the Million Women Study some 20 years ago and approved as MREC 97/5/001 (Appendix 1). The Million Women Study is an open-ended prospective study of the health of women in England and Scotland, and includes 1.3 million women recruited between 1996 and 2001 at age 50-64 years through NHS Breast Screening Centres [1]. The cohort has continuing follow-up through postal and online re-surveys, to update exposure measures, and through electronic linkage to routinely-collected NHS records for disease outcomes, including death, cancer registration and hospital admissions databases.

The background to the study was the rapidly increasing use of HT in the 1990s in the UK and elsewhere, stimulated in part by claims that use of HT use could improve general well-being and increase life expectancy. The worldwide evidence was beginning to show, however, that HT preparations increased breast cancer risk, and there was little information about the effects of the type of HT most commonly used in Europe, containing both oestrogens and progestogens [2]. It was also clear that women born in the 1940s, who reached adulthood in the 1960s, had considerably different lifestyles to previous generations. For example, large proportions had begun using oral contraceptives and smoking as teenagers, and the long-term effects of these behaviours had not been well studied in women. At the same time there was growing concern during the 1990s about the increasing prevalence of obesity, and claims that other factors such as diet had important effects on health, all of which required large scale prospective evidence to study reliably.

The co-ordinating centre for the Million Women Study is based in the Cancer Epidemiology Unit, Nuffield Department of Population Health, University of Oxford. The study was set up in collaboration with the NHS Breast Screening Programme, and is funded mainly by the UK Medical Research Council and Cancer Research UK. Full information, including copies of study questionnaires and mechanisms for data access, can be found on the study website (www.millionwomenstudy.org). The study is run by the Principal Investigators, Professor Dame Valerie Beral, Professor Jane Green and Professor Gillian Reeves with advice from the Million Women Study Advisory Committee (Appendix 4).

5. AIMS AND OBJECTIVES

The initial stimulus for the study was to obtain robust prospective information on the risk of breast cancer associated with use of different types of menopausal hormone therapy (HT). When planning the necessary large-scale prospective study, an equally important aim was to obtain reliable information on the effects of other potentially modifiable factors that affect women's health in middle age and in old age.

First study results for hormone therapy and breast cancer were published in 2003 [3], showing that use of oestrogen-progestogen preparations causes much greater increases in the risk of breast cancer than oestrogen-only preparations; further analyses described the overall impact of HT use on risk for breast, ovarian and endometrial cancers [4,5]. We have since published the results of investigations of risk for vascular, musculoskeletal and neurodegenerative disease, as well as for various cancers, in relation to smoking, obesity, alcohol consumption, body size and other factors (all publications are listed on the study website).

The mature cohort represents a unique health research resource with now over 20 million person-years of follow-up (some 15 years on average per woman), with detailed individual information on behavioural, anthropometric and demographic exposures over many years and with virtually complete unbiased follow-up through electronic linkage to NHS databases. We have published about 100 peer-reviewed papers. We continue to aim to provide reliable evidence to inform individual and public health decisions, focussing on modifiable risk factors for common and serious diseases of middle and older age in women. An important part of our work is in collaboration with other researchers and other resources, to maximise the benefit for public health of the many years of investment by participants, study investigators and funders.

6. STUDY DESIGN

The Million Women Study is an open-ended ongoing prospective cohort study.

Ethics approval for the recruitment of the study population was obtained initially in 1996 for individual regions and as a multi-centre cohort study in 1997 (Appendix 1). A total of 31 substantial amendments to the original protocol have been approved to date (Appendix 2). These cover content and mailing for additional (eg re-survey) questionnaires, and accompanying leaflets describing main study findings; new data linkages; content and introduction of online questionnaires; and validation studies.

6.1 Recruitment and participants

Women were recruited between 1996 and 2001 through the NHS Breast Screening Programme, which invites all UK women of a specified age for free routine breast screening every three years. In 1996-2001 the programme routinely invited women aged 50-64 years, by sending each individual a letter offering them a specific date and time at a specific screening centre for their mammographic screen. The administrative structure of the screening programme in 66 of its screening centres (about half of the breast screening centres in the UK) permitted us to include the Million Women Study questionnaire in the invitation letter for screening. Pilot studies in 1994-1996 had shown that inclusion of a questionnaire with the invitation did not affect uptake of breast screening [6].

Women eligible to join the study were all those registered with an NHS General Practitioner in areas covered by the 66 participating breast screening units, and aged 50-64 years, and so sent routine invitations for breast screening. These women received a study questionnaire with their screening invitation and were asked, if interested in joining the study, to bring the completed questionnaire (including consent form) with them when they attended for screening. The eligible women included about half of the UK population of women aged 50-64 years. About half the women invited to join the study brought completed questionnaires with them when they came for screening, or posted completed questionnaires to the Million Women Study co-ordinating centre. Thus the Million Women Study recruited 1 in 4 of the entire UK population aged 50-64 years in 1996-2001. The recruitment questionnaire asked about women's health, lifestyle, reproductive history, HT use, and basic demographic and anthropometric information. Women joining the study consented to re-contact and to follow-up through their medical records, on the understanding that we would keep their data safe

and use them only for medical research. Characteristics of the women at recruitment are shown in *Table 1, Appendix 5*.

6.2. Participant withdrawal

Women are free to withdraw from the study at any time, and this has been stated (together with postal and Freephone contact details) on postal communications sent to all participants (questionnaires, participant information sheets, leaflets) throughout the study. Full details of the options for withdrawal are on the study website. Women who contact the study asking to withdraw are offered the options of (1) receiving no further direct contact from us (by post or email), but allowing us continued use of the data they have already provided, and continued follow-up through electronic linkage; (2) no further contact and no further follow-up, but continued use of data already collected; (3) full withdrawal: no further contact, no further follow-up and no new use of collected data.

Mental Capacity

When we are informed that a participant no longer has mental capacity to complete our questionnaires (usually by a family member, following a re-survey mailing), we withdraw the participant from all further contact. We write to the informant advising them of this and asking for their agreement to continue to use the information already provided by the participant, and to continue routine follow-up through data linkage. If they object, we implement full withdrawal. Our process complies with the Mental Capacity Act (2005).

6.3 Follow-up

Re-survey questionnaires

All surviving women in the study have been contacted every 3-5 years since recruitment, and asked to complete postal re-survey questionnaires to update information on behavioural factors- such as smoking, weight, diet, use of medication - which may have changed (and for which information is not readily available from other sources). In addition, re-surveys provide the opportunity to ask additional questions about new hypotheses and other issues which become increasingly relevant as the cohort ages and new health conditions become more common. Some 80% of women in the cohort have responded to at least one of the four re-survey questionnaires. Since 2010, over 100,000 women have also completed online questionnaires to supplement information on diet and physical activity, and we plan further online re-surveys. Details of dates and content of the re-survey questionnaires are available on the study website and are summarised as *Table 2, appendix 5*.

Biological samples and physical measurements

In 2004 separate ethical approval was obtained to collect biological samples from women in the cohort (Million Women Study: Disease Susceptibility in Women, REC 03/5/071; protocol is Appendix 3). Some 60,000 women have provided blood samples, collected for us by their General Practitioners and Practice staff. These allow us to add genetic and biochemical analyses to our cohort [7,8]. As part of the Disease Susceptibility in Women study we also collected physical measurements (height, weight, waist and hip circumference, blood pressure) from some 4000 women in the cohort for validation of self-reported values.

Linked outcomes data

For follow-up of outcomes, all women in the study have been 'flagged' on the NHS Central Registers for regular follow-up through data linkage for deaths and cancer registrations (and for information on changes in registration with the NHS, eg through emigration) since recruitment; and since 2005/6 also for hospital (day case and in-patient) admissions. For subsets of the cohort, linkage has been established also to other data sources, including cancer screening programme and primary care databases (see *Table 3, Appendix 5* for details of data sources and linkage, and *Table 4, Appendix 5* for selected outcome numbers). Availability of routinely-collected NHS data is increasing, and we hope to be able to take advantage of this to enhance the cohort in future.

We have conducted a number of studies to confirm the reliability of the routinely-collected outcomes data for use in risk factor analyses, including comparisons with hospital medical notes and with both linked coded and directly-obtained information from primary care [9,10,11].

6.4 Future plans

We plan to continue to contact surviving participants through post or email to maintain up to date information on major health behavioural and social factors (eg diet, weight, smoking, medication use, caring responsibilities, social contact). We will continue to develop new data linkages as datasets become available.

Our currently funded work programme (2012/13 to 2018/19) includes work on risk factors for breast, prostate, colorectal, ovarian, endometrial and less common cancers, with an emphasis on identifying risk factors for different tumour subtypes, and on contributing to international collaborative analyses of the worldwide evidence on hormonal factors. Non-cancer work includes investigation of potentially modifiable risk factors (use of HT and other drugs, and behavioural health factors such as physical activity) for osteoporotic fracture, dementia and vascular disease. In addition to work quantifying the effects on women's health of long-term smoking, body size throughout life and diet and medications, we are looking at disease risk in relation to social factors, disability and frailty, and ethnicity.

Research collaboration
An important aspect of our work is in collaboration with other research studies, in particular with other similar cohorts in the UK and elsewhere, to minimise duplication of effort and to ensure that important health questions are answered by the studies best placed to do so. Details of our study are listed on the Medical Research Council's Cohort Directory. The design and development of our web-based questionnaires has been conducted jointly with UK Biobank and many of our current and future analyses will use data from both cohorts. We are a contributing cohort for the Dementias Platform UK, and long-standing contributors to international collaborations on hormonal factors in breast, cervix, ovarian and endometrial cancers. In Oxford, the study's established position as one of several large prospective studies within the Nuffield Department of Population Health (others include the China Kadoorie Biobank, studies in Mexico, Cuba and India, and EPIC-Oxford) offers opportunities for collaboration with other population health groups (eg health economics) and close links to a new Big Data Institute. We work together with clinical groups in Oxford, and with the UK NHS Cancer Screening Programmes.

6.5 End of study
We plan to continue the study for as long as is both practicable and useful. No set end date has been decided. Data linkage now allows very long term, cost-effective follow-up, of high scientific value, and we expect that the study will remain active for several more years. Women in the study were born in the 1930s and 1940s and are now reaching the ages at which conditions such as dementia become common enough to study reliably. Study progress and value are regularly reviewed by the University and by funders.

7. STATISTICAL METHODS AND ANALYSIS

With very large numbers and very little missing data, most analyses to date have used standard statistical techniques. Analyses of the risk of disease outcome in relation to categorical or continuous measures of exposure generally use Cox proportional hazards regression to estimate Hazard Ratios adjusted for potential confounders. In studies of genetic risk factors, where exposure information was obtained on a nested case-control series, logistic regression models are used to obtain odds ratios of disease according to genotype. Analyses of risk factors for disease according to disease subtype are carried out using a competing risks approach and formal tests for heterogeneity of association by subtype are assessed using standard chi-squared statistics. Where appropriate, changes in exposure during follow-up (eg menopausal status, HT use) are taken into account by using time dependent covariates. For those anthropometric variables where physical measurements are available on a subset of the cohort (eg weight, height), trends in disease risk are adjusted to take account of possible measurement error. Where appropriate, the potential impact of reverse causality on the relationship between risk factors and disease outcome is assessed by stratification of analyses according to health status or time under follow-up.

For relatively common outcomes (eg breast cancer, cardiovascular disease, musculoskeletal disease) the study has substantial power to detect even modest associations, and to study associations by relevant subtypes of disease. For comparatively rare diseases (eg cancers of the oesophagus, bladder and pancreas), there is still sufficient power to detect a relative risk of at least 1.2 associated with

extreme quartiles of a given exposure. Some questions are difficult to reliably address within a single study and in such cases, we contribute anonymised data to relevant international collaborative groups with the aim of leveraging greater statistical power and putting study-specific findings in a wider context.

8. DATA MANAGEMENT

8.1 Data Access

The study welcomes requests from bona-fide researchers for access to study data for high-quality medical and health research, and our data access policy (Appendix 6) fulfils Research Councils UK requirements. We follow a controlled access data sharing model: our commitment to our participants, the need to safeguard the reputation of this ongoing research study and restrictions on third-party access to linked NHS data mean that sharing of data is often through collaboration. As was usual at the time, the issue of data sharing was not raised when women gave consent for follow-up at study recruitment and so we do not have explicit consent for sharing data with external researchers.

Requests for data and for collaboration are considered by the study PIs and Advisory Committee with oversight by the Data Access Oversight Committee of the NDPH Richard Doll Centenary Archive, which provides a coordinated structure for data sharing within the Department (see Appendix 6 and https://www.ndph.ox.ac.uk/about/data-access-policy). Study data provided to external researchers is de-identified and labelled with a study-specific pseudo-identifier. Under our current Data Sharing Agreements, linked individual data provided by ONS, NHS Digital, PHE or any other provider are not released to external researchers.

Direct access will be granted to authorised representatives from the Sponsor and host institution for monitoring and/or audit of the study to ensure compliance with regulations.

8.2 Data Recording, Storage, Record Keeping and Data Confidentiality

Data security and governance

Million Women Study records are held securely within the Cancer Epidemiology Unit of the Nuffield Department of Population Health. The University of Oxford acts as Sponsor and Data Controller. Data security complies with the Data Protection Act (University of Oxford: registration number Z575783X) and with Unit, Departmental and University data security policies. The Cancer Epidemiology Unit has current NHS Information Governance Toolkit accreditation for storage of linked NHS data (ref. 8J229; relevant data security and governance policies are available on the Cancer Epidemiology Unit website, https://www.ceu.ox.ac.uk/policies2). Questionnaire, blood sample and linked follow-up data are stored separately with restricted access within the study team, and used to prepare de-identified datasets for analysis. Identifying information provided by participants (name, address, date of birth, email addresses) and through NHS central registers (current address and registration status with NHS) is held by the study for the purposes of re-contact and data linkage. Data provided for linkage or to external researchers is labelled with project-specific pseudo-identifiers to minimise risk of participant re-identification. All those receiving data from the study are also bound by legal agreement not to identify participants, and where possible data are provided in a form which also minimises this risk (eg as tabulations, or suppressing part of birth date, or small numbers in data cells).

Physical security

CEU resides in a secure building with swipe card access on all external doors. External doors are monitored by CCTV. Visitors and deliveries are required to report to the Reception for verification by Reception staff. High security areas (e.g. server rooms) are physically and electrically separate from other CEU facilities and have additional security locks in place. Access is restricted to relevant staff. Server rooms have air conditioning units to ensure that the servers operate within operating limits specified by the equipment manufacturers. Offices are secured by door locks out of normal working hours and when not in use. Server room power is supplied from multiple mains feeds, with equipment split between feeds. Main servers and other key hardware are protected by uninterruptible power supply units (UPS) in order to maintain service in the event of a power outage and prevent corruption of information.

Paper records (including consent forms and completed study questionnaires) are stored in

locked cabinets/rooms within the Richard Doll Building. As soon as is practicable, paper documents are scanned to electronic files for long term storage and the paper copies shredded and disposed of as confidential waste.

Back-up tapes
Routine daily backup tapes are stored on-site in a fire safe. Archive tapes are stored at a secure off-site location (3-6 month rotation).

Electronic security
All system level utilities are separated from application programs and available to authorised users only by a minimum of separate user ID/login procedures. Some systems only allow access at this level from the system console which will be in a secure room. Servers will have unused facilities removed before use. Terminals unused for a given length of time will time-out to a secure login. No terminal facilities will be available in areas accessible to the general public (i.e. outside of the CEU controlled areas). Root access to servers will normally be limited to IT Support staff only with individual access identified through an initial individual user login. Root user access will only be used for initial configuration, backups, emergency access to servers and configuration changes (where required). All root level logins will be audited in the system logs.

8.3 Data Linkage
Million Women Study participants are identified by 'flags' on the NHS Central Registers and for England and Scotland and no further transfer of identifying information is required for continuing follow-up through Office for National Statistics or NHS records. Death, cancer registration and hospital admissions data are provided by NHS Digital (the Health and Social Care Information Centre, HSCIC) in England and by the Information Services Division of the NHS in Scotland (ISDScotland). Additional data linkages have been obtained through Public Health England (NHS Cancer Screening Programmes; Cancer Services and Outcomes data for additional details of registered cancers) and through the Clinical Practice Research Datalink for primary care data in England.

For future approved data linkages, identifiable information supplied by the Million Women Study may include name, NHS Number, date of birth and postcode and will be subject to the Unit's policies on secure transfer of data.

9. ETHICAL CONSIDERATIONS

The overriding ethical considerations for the study are

- to fulfil our promise to study participants to use their data only for medical research, and to keep the data safe
- to use study data in such a way as to maximise their value for public health

There are no evident risks to participants from taking part in the study.

9.1 Consent and fair processing
Million Women Study participants all provided signed consent to re-contact and to follow-up through medical records when they joined the study in 1996-2001. Women providing blood samples (from 2006 onwards) gave separate informed consent for the immediate and future use of their samples for genetic and biochemical research.

As noted above, all surviving women in the cohort have been contacted directly by post with re-survey questionnaires (and accompanying leaflets or letters describing study progress) every 3-5 years since recruitment, most recently in 2013-2014. Some 80% of women in the study have responded to at least one of these contacts. Full information on the study, including regularly-updated progress reports and all study publications, has been available on the study website (set up in 2000), and is provided on request to women who do not have internet access. The study provides a Freephone line (and later also an email address) for enquiries and individual replies are sent to participants (and non-participants) who contact us.

As noted in section 6.4, we plan continued contact with surviving participants to update health and social data, and will use such opportunities to keep participants informed of the evolving study

purpose, the data being linked and the data providers involved (eg ONS, NHS Digital, PHE, CPRD, ISD Scotland).

9.2 Loss of mental capacity during follow-up
As noted in section 6.1, where we are informed of loss of mental capacity of a participant we take steps to ascertain the wishes of the family member informing us about continued use of data and/or follow-up.

9.3 Discontinuation/withdrawal of participants from the study
Women are free to withdraw from the study at any time. Details of how to withdraw, and of the options available (no further contact; no further follow-up through data linkage; no further use of data) are described on the study website, and re-survey mailings sent to all participants include contact details for withdrawal.

9.4 Declaration of Helsinki and other regulations
The Investigators will ensure that this study is conducted in accordance with the principles of the Declaration of Helsinki and in accordance with relevant guidelines and regulations.

9.5 Guidelines for Good Clinical Practice
The University will ensure that this study is conducted in accordance with relevant regulations and Good Clinical Practice.

9.6 Approvals
This protocol will be submitted to the Sponsor (University of Oxford) for written approval and to the East of England, Cambridge South Research Ethics Committee as a substantial amendment to the Million Women Study.

The PIs submit once a year throughout the study or on request, an annual progress report to the REC Committee and the Sponsor. The investigators will submit, and where necessary, obtain approval from the above parties for all substantial amendments to the study protocol or to participant information, consent and contact documents. In addition, an end of study notification and final report will be submitted to the same parties.

9.7 Participant and Data Confidentiality
The study staff will ensure that the participants' anonymity is maintained. All documents will be stored securely and only accessible by study staff and authorised personnel. The study will comply with the Data Protection Act, which requires data to be anonymised as soon as it is practical to do so. The study holds identifying data on participants for the purposes of re-contact and study follow-up, but where possible study data files are pseudo-anonymised using a participant ID number.

9.8 Expenses and Benefits
There are no past or intended payments to participants or any other benefits.

10. IMPACT OF THE STUDY

The strength of this study is, as described above, its ability to provide reliable answers to questions of public health importance. As a mature, very large-scale observational study with prospective collection of exposure data and virtually complete, unbiased long term follow-up through linkage to routinely collected NHS data, it is ideally placed to investigate many outstanding questions of public health importance, both within the cohort and through collaboration with other research resources. The conditions under study are both common and serious and we focus on risk factors, such as diet, smoking, obesity and physical activity which are largely modifiable, offering opportunities for disease prevention. The observational nature of the study allows flexibility in studying new outcomes and exposures as they become of public health interest; and as an established cohort, continued follow-up through data linkage is both feasible and cost-effective.

The value of the study in terms of public health benefit has been proven, most notably in its contribution to evidence-based guidelines for the prescribing and use of HRT, which have led to a fall in inappropriate HT use and subsequently to falling rates of breast cancer (refs). The study has also

contributed the first reliable evidence on effects of long-term smoking in women (and of quitting smoking), new evidence suggesting little or no impact of night shift work on breast cancer risk, and the first comprehensive assessment of the magnitude of the risk of venous thromboembolism after surgery.

We place a high value on wide dissemination of our results, through peer-reviewed publications and through personal interactions with the media, the public, the UK and international research community and health policymakers.

11. FINANCE AND INSURANCE

The Million Women Study is supported by grants from Cancer Research UK (current grant no. C570/A11692) and the Medical Research Council (current grant no. MR/K02700X/1).

The University has a specialist insurance policy in place which would operate in the event of any participant suffering harm as a result of their involvement in the research (Newline Underwriting Management Ltd, at Lloyd's of London).

12: PUBLICATION POLICY

The Investigators will be involved in reviewing drafts of the manuscripts, abstracts, press releases and any other publications arising from the study. Authors will acknowledge study funding appropriately. Authorship will be determined in accordance with the ICMJE guidelines and other contributors will be acknowledged. University policy on Open Access publication will be followed, and all publications are listed on the study and Cancer Epidemiology Unit websites.

13. REFERENCES

1. Banks E, Beral V, Reeves G on behalf of the Million Women Study Collaborative Group. The Million Women Study: Design and Characteristics of the Study Population. Breast Cancer Research 1999, 1:73-80.

2. Collaborative Group on Hormonal Factors in Breast Cancer. 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 1997; 350:1047-1059.

3. Beral V, Banks E, Reeves G, Bull D on behalf of the Million Women Study Collaborators Breast cancer and hormone-replacement therapy in the Million Women Study. Lancet 2003; 362:419-427.

4. Beral V, Bull D, Reeves G, for the Million Women Study Collaborators. Endometrial cancer and hormone-replacement therapy in the Million Women Study. Lancet 2005; 365:1543-51.

5. Beral V, Bull D, Green J, Reeves G on behalf of the Million Women Study Collaborators. Ovarian cancer and hormone replacement therapy in the Million Women Study. Lancet 2007; 369:1703-1710.

6. Banks E, Richardson A, Beral V, Crossley B, Simmonds M, Hilton E, English R, Davis J, Austoker J. Effect on attendance at breast cancer screening of adding a self-administered questionnaire to the usual invitation to breast screening in Southern England. J Epidemiol Community Health 1998; 52:116-119.

7. Travis RC, Reeves GK, Green J, Bull D, Tipper SJ, Baker K, Beral V, Peto R, Bell J, Zelenika D, Lathrop M. Gene-environment interactions in 7610 women with breast cancer: prospective evidence from the Million Women Study. Lancet 2010: 375(9732): 2143–2151.

8. Reeves GK, Travis RC, Green J, Bull D, Tipper SJ, Baker K, Beral V, Peto R, Bell J, Zelenika D, Lathrop M. Incidence of Breast Cancer and Its Subtypes in Relation to Individual and Multiple Low-Penetrance Genetic Susceptibility Loci. JAMA 2010;304: 426-434.

9. Wright FL, Green J, Canoy D, Cairns BJ, Balkwill A, Beral V; Million Women Study Collaborators. Vascular disease in women: comparison of diagnoses in hospital episode statistics and general practice records in England. BMC Med Res Methodol. 2012 Oct 23; 12:161. doi: 10.1186/1471-2288-12-161.

10. Wright L et al. Validity over time of self-reported anthropometric variables during follow-up of a large cohort of UK women. BMC Med Res Methodol. 2015 Oct 8;15:81.

11. Brown A, Kirichek O, Balkwill A, Reeves G, Beral V, Sudlow C, Gallacher J, Green J. Comparison of dementia recorded in routinely collected hospital admission data in England with dementia recorded in primary care. Emerg Themes Epidemiol. 2016 Oct 28;13:11. PubMed PMID: 27800007.

14. APPENDICES

APPENDIX 1 Original Million Women Study Protocol, as approved REC 97/5/001

THE MILLION WOMEN STUDY

a national survey of women invited for breast screening

PROTOCOL FOR THE STUDY

The principal aim of this research is to obtain reliable information on the risk of breast cancer associated with the use of hormone replacement therapy (HRT). Recent evidence suggests that the risk may be increased, but insufficient data are available in the entire world literature to describe that risk accurately or to determine whether HRT preparations containing oestrogen alone or oestrogen and progestogen have different effects. Large numbers of women, i.e. more than 5,000 with breast cancer, need to be studied to answer the questions reliably, and the NHSBSP offers a unique opportunity to do so. Indeed it would probably be impossible to carry out a study of the scale required anywhere else in the world. An additional consideration is that the age range of women who use HRT corresponds with the age when screening is offered, and a subsidiary question that would be answered by the proposed study is whether HRT use affects the sensitivity and specificity of mammography.

Summary of the study

The proposed study would involve sending a brief questionnaire to women when they are invited for screening. The questionnaire asks about factors known to affect breast cancer risk and about the use of hormone replacement therapy and other hormones. It also asks women for permission to obtain follow-up information about their health, and to provide personal details. Follow-up for cancers detected at mammography would be via the screening clinics and for other cancers and deaths via the NHS Central Registers. The risk of cancer and of death from various causes in women who had used HRT will then be compared to the risks in women who had not used HRT.

Results from pilot studies

Extensive piloting of the study has already been carried out on over 6000 women in West London and Oxford. These have demonstrated that the questionnaire is acceptable to women: a sample of women who were sent the questionnaire were interviewed by telephone and none expressed concern about any of the questions asked. By far the most common questions were about the effects of HRT, and many women said how pleased they were that this research was being done. Pilot studies have also shown that attendance rates did not differ significantly in women who were and were not sent a questionnaire. Furthermore, among the women who attended for screening, 78% completed the questionnaire, and 95% of these women gave permission for follow-up.

In conclusion, the results from pilot studies demonstrate that the enclosed questionnaire is acceptable and that sending it out to women does not alter attendance rates for screening. The

response rate is high and women seem keen to take part in this type of research. Additionally, the prevalence of use of HRT is high, with 41% having used HRT at some time and 28% being current users of HRT.

Details of the study

Eligibility: All screening centres in the NHSBSP will be eligible to collaborate in this research and all women invited for screening will be eligible to take part. Participation is entirely voluntary.

Approach to clinics: Each breast screening centre will be approached and asked if it wishes to collaborate in this research. If so, the study team will discuss how best to arrange the sending out and collection of questionnaires and seek local ethical committee approval for the study.

The study is already underway in about 40 screening centres in England and the experience so far is that each centre has its own particular routine: an important principle of the study is that it should interfere as little as possible with the normal running of the screening programme. Therefore, there will be regular communication between the study team and staff at each participating centre to ensure that involvement in the study causes as little extra work as possible and that problems that arise are dealt with swiftly.

Approach to subjects: Women will be sent the attached questionnaire at the time they are invited for screening and asked to bring it with them when they attend for screening. A freephone number is provided if the women have any questions.

Follow-up: Once each year routine information about the breast cancers diagnosed at screening will be sought from each screening clinic. In addition, the study team will seek information about cancer incidence and deaths from the NHS Central Registries.

Logistics: Staff at the Million Women Study Co-ordinating Centre , Oxford, will be responsible for the preparation and delivery of the questionnaires to the screening centres, for the collection of the questionnaires **and** for the follow-up. The screening clinics will be responsible for despatching the questionnaires to the women and for providing, once each year, routine details about the cancers diagnosed in study participants.

Numbers: At least 5,000 women with breast cancer are needed to detect a 25% difference in breast cancer risk in long term users of oestrogen alone compared with users of combinations of oestrogen and progestogen. In addition at least 500 interval cancers are needed in the first year after screening to see whether HRT use affects sensitivity of screening. To obtain such numbers about 1,000,000 women will need to be recruited.

Timetable: It is hoped that most centres that wish to participate will have begun sending questionnaires out by late 1997. Recruitment will last for about 3 years in each centre. The exact timing will be determined by the number of centres that collaborate.

Confidentiality and ethics: Local Ethical Committee approval will be sought separately for each screening centre. Participation in the study will be entirely voluntary. It will be made clear to women that they do not have to take part in the study if they do not want to and that their decision does not affect their management. All information provided will be stored in accordance with the regulations of the Data Protection Act (registration number with the Office of the Data Protection Registrar: K3039784). The data will be treated with utmost

confidentiality and used only for medical research. Only the study team will have access to computerised data, via passwords. Any publication resulting from this work will not identify the individual women who took part.

Publication of results: This is a joint research project of the NHS Breast Screening Programme and the Imperial Cancer Research Fund (ICRF). Publications of the main results will be in the name of "The Million Women Study Group", and all collaborators will be named.

Appendix 2: List of approved amendments to Million Women Study protocol, to August 2017

Amendment No.	**Protocol Version No.**	**Date**	**Changes made**
Amendment 1	Version 1 May 1997	Approved 2Dec1998	**3-year re-survey postal questionnaire** Re-survey(yellow) questionnaire to be sent to the study population
Amendment 2		Approved 8Dec1999	**Continuation of study and revision of 3-year re-survey questionnaire.** Part 1 unlimited approval for study granted for the continuation of the study Part 2 revisions to 3-year questionnaire
Amendment 3		Approved 8Mar2000	**Additional postal questionnaires to obtain further details** Additional questionnaire for women who reported at the 3-year re-survey that they have been diagnosed with breast cancer. Additional questionnaire for women who reported at the 3-year re-survey that they had an abnormal cervical smear.
Amendment 4		Approved 19Jan2006	**8-year re-survey postal questionnaire** 8-year resurvey questionnaire to be sent to survivors
Amendment 5 (AM01)		Approved 15Jan2008	**Data linkage for body size measure validation** Approval to link to corresponding data for women who also participated in the MRC 1946 British Cohort Study
Amendment 6 (REC amendment 2-designated Am 5)		Approved 10Jan2009	**Validation of self-reported questionnaire data on vascular disease by GP review of primary care medical records** This included a letter to GPs explaining the aim of the validation study and asking for confirmation of diagnoses of ischemic heart disease, stroke and venous thromboembolism
Amendment 7 (REC amendment 3)		Approved 02Oct2009	**12-year re-survey postal questionnaire** 12-year resurvey questionnaire to be sent to survivors, information leaflet and reminder leaflet
Amendment 8 (REC amendment 4)		Approved 10Nov2009	**Minor change to wording of 12-year re-survey questionnaire**
Amendment 9 (REC amendment 5)		Approved 13Jan2010	**Minor changes to questionnaire, information leaflet and reminder leaflet for the 12-year re-survey**
Amendment 10 (REC		Approved 08Feb2010	**Introduction of on-line 24-hour recall diet questionnaire**

amendment 6)			On line diet questionnaire to be sent to women who have provided an email address at the 12-year resurvey (Amend 7)
Amendment 11 (REC amendment 7)		Approved 23Feb2010	**Addition of consultee letter to inform those participants' representatives who have given notice of a participants' mental incapacity of their options under the Mental Capacity Act 2005,** The approval of these documents enables us to write to relatives or representatives of participants who have suffered a mental incapacity since the last time they were contacted for the study. The representatives will receive a letter outlining options, for continuing to use existing data we hold for the participant on the study, offering them the option of writing to Prof Beral to withdraw data already held for the participant. This complies with legislation under the Mental Capacity Act (2005)
Amendment 12 (REC amendment 8)		Approved 05March2010	**Revisions of wording for vascular disease validation study** Inclusion of a letter to Practice Managers to maximise the response rate; and addition of coding boxes to reply forms
Amendment 13 (REC amendment 9)		Approved 17June2010	**Minor changes to wording of online 24 hour recall diet questionnaire, together with disabling of PIN after five failed login attempts**
Amendment 14 (REC amendment 10)		Approved 07July2010	**Minor wording revisions for letters used in Vascular Validation Study** Change of contact person at co-ordinating centre and of payment arrangement details on Practice Manager letter
Amendment 16		Approved 11Aug2011	**Changes to the 12-year resurvey questionnaire for the remaining mailings**
Amendment 17		Approved 19Sept2011	**Validation of hospital diagnoses of neurodegenerative disease.** GP letter asking for confirmation of diagnosis of Motor Neuron disease, Alzheimer's disease, dementia, Parkinson's disease
Amendment 18		Approved 19 Jan 2012	**Changes to online 24 hour recall diet questionnaire** Participants to receive emails once a year instead of every 3 months (REC confirmed this was a non-substantial amendment)
Amendment 19		Approved 02.May.2012	**Researchers to contact participants GPs for information on prescribing of Bisphosphonates**
Amendment 20 (labelled as number 17 by NRES)		Approved 10 July 2012	**Request to link data from Million Women Study with data held by UK Biobank**
Amendment 21		Approved 27 Sept 2012	**Re-contact participants about 24 hour recall online diet questionnaire with minor changes to two emails.** Minor corrections to documents acknowledged by sponsor
Amendment 22		Approved 11 Oct 2012	**New online food frequency questionnaire (FFQ)**

			Request to email all participants in study via email, to invite them to complete new online food frequency questionnaire.
Amendment 23 non-sub		Acknowledged 23Jan 2013	**Correction of minor typographical errors on FFQ questionnaire**
Amendment 24		Approved 26 Feb 2013	**15-year re-survey questionnaire** 15-year resurvey questionnaire to be sent to survivors
Amendment 25		Approved 16 May 2013	**Minor revisions to 12-year resurvey questionnaire**
Amendment 26		Approved 30May 2014	**Minor changes to the contents of the online 24 hour recall diet questionnaire, based on feedback from participants**
Amendment 27	Version 1.1 March 2015	Approved 01 April 2015	**Pilot study for processing tumour tissue samples of participants with breast cancer treated at Oxford University Hospitals NHS Trust**
Amendment 28		Approval 04 March 2016	**Processing tumour tissue samples of participants with ovarian cancer treated at Oxford University Hospitals NHS Trust**
Amendment 29	Version 1.2 March 2016	Approval 04 April 2016	**Validation studies comparing the linked outcomes data with detailed clinical records held at the Oxford University Hospitals NHS Foundation Trust (OUH) in Oxford**
Amendment 30	Version 1.3 April 2016	Approved 12th May 2016	**To add additional questions to the FFQ online diet questionnaire and invite participants who provided an email address to complete it.**
Amendment 31	V1.4 Sept 2016	Approved 31 October 2016	**Minor changes to the contents of the emails sent out to our participants about online diet questionnaires**
Amendment 32	V2.0	Approved 07 March 2017	**Updated protocol**
Amendment 33	V2.1	Approved 11 September 2017	**Online cognitive function questionnaire.** Invitation to participants to complete a short assessment online. Similar methods to those used for the first and second online diet questionnaire (Amendments 10,22 and 31 approved)
Amendment 34	V2.2		**Fifth general follow-up questionnaire.** All women active in the study (~1.2 million) will be eligible for the 5th follow-up mailing, i.e. all women who were recruited to the study, except those who have since died, left the NHS or have asked us not to contact them again.

Appendix 3: Approved protocol for collection of blood samples in Million Women Study participants (REC 03/5/071, 2004) and list of subsequent amendments.

THE MILLION WOMEN STUDY

THE MILLION WOMEN STUDY: DETAILED INVESTIGATION OF SUSCEPTIBILITY TO DISEASE

Short title: The Million Women Study: Disease Susceptibility in Women

Aim
To investigate in detail factors affecting susceptibility to disease within the Million Women Study, including the separate and joint effects of exposure to environmental factors, such as hormone replacement therapy, and of the individual's genetic constitution. The main focus will be on susceptibility to breast cancer and to cardiovascular disease.

Background
The Million Women Study (Eastern MREC 97/5/01; see Flow Chart 1) is a nationwide prospective cohort study set up primarily to investigate the effects of different types of hormone replacement therapy (HRT) on the risk of breast cancer. Over 1.3 million women aged 50-64 (1 in 4 of UK women in this age group) were recruited through NHS breast screening centres between 1996 and 2001. Prevalent and incident disease is monitored through self-reporting on recruitment and follow-up questionnaires and by linkage to the National Health Service Breast Screening Programme, Cancer Registries and the Office of National Statistics. The large scale of the Million Women Study and the detailed prospective information collected on environmental risk factors such as the use of HRT mean that it offers an unprecedented opportunity to study the separate and joint effects of both environmental and genetic risk factors for disease.

Rationale
Breast Cancer
Breast cancer is the commonest cancer among women in the UK and is an important cause of morbidity and mortality. There is considerable public and scientific interest in the role of genetic factors in susceptibility to breast cancer. While there is mounting evidence on the substantial risk of breast cancer associated with rare high-risk gene mutations such as *BRCA1* and *BRCA2*, much less is known about the possible effect on breast cancer risk of more common lower-risk variations (polymorphisms) in genes such as those (eg *CYP17, CYP19, HSD17B1*) involved in the metabolism of sex hormones like oestrogen. Although the risks of breast cancer associated with these lower-risk gene variants are likely to be relatively modest, they are potentially of considerable public health importance as the genetic variations involved are very common (present, on average, in about 20-60% of the population in the UK). Particularly important are the possible joint effects on breast cancer risk of these common genetic variants and equally common environmental risk factors such as the use of hormone replacement therapy (at recruitment a third of women in the Million Women Study were current users of HRT). These joint effects can only be adequately investigated in very large studies. Initial studies in this field have been limited by small study size, overemphasis on marginal findings and failure to replicate positive results; however it is hoped that relevant polymorphisms will be more reliably identified in larger-scale studies currently under way. The Million Women Study has accrued 10 000 incident cases of breast cancer within the first three years of follow-up and over 35 000 incident cases are expected by 2008. We plan to focus our breast cancer susceptibility studies on the investigation of strong candidate polymorphisms once these have been

identified, and, unlike previous studies, we will be able to take into account detailed information on tumour histology and on oestrogen and progesterone receptor status as well as prospective data on reproductive factors and HRT use.

Cardiovascular disease

Current users of HRT are at approximately three-fold increased risk of venous thromboembolism (deep venous thrombosis and pulmonary embolism). The effect of HRT on the risk of arterial disease, including coronary heart disease and stroke, is less clear; recent results from randomised trials suggest that women using HRT may be at increased risk of stroke (relative risk 1.3) and that the risk of coronary heart disease may be increased, at least in the first year of HRT use. These results contrast with the apparent protective effect of HRT on risk of cardiovascular disease seen in earlier observational studies (and which may have been due to prescribing bias). Venous and arterial disease are common and important causes of morbidity and mortality in post-menopausal women. Genetic factors such as polymorphisms in the genes involved in blood clotting (eg Factor V Leiden and prothrombin G20210A) are known to increase the risk of venous thromboembolism, and may also affect the risk of arterial thrombosis. Evidence from oral contraceptive use suggests that the joint effect of such thrombophilic mutations and exogenous hormone use may be substantial. Compared to non-carriers of the Factor V Leiden mutation who were not using oral contraceptives, the relative risk for venous thromboembolism was 4 for non-carriers using oral contraceptives, 8 for carriers not using oral contraceptives and 35 for carriers using oral contraceptives. In the only published study of polymorphisms of prothrombotic genes and HRT use the risk of venous thromboembolism was increased independently by HRT (relative risk 3) and by Factor V Leiden (relative risk 4); the relative risk in women with both factors was 15. HRT use and thrombophilic mutations are both relatively common and the Million Women Study provides the opportunity to investigate in detail their separate and joint effects on cardiovascular disease risk. It is possible that consideration should be given to screening women for markers of susceptibility before HRT is prescribed; about 5% of women carry the Factor V Leiden mutation, one in thirty carriers of Factor V Leiden may develop venous thromboembolism each year on HRT and this susceptible subgroup of carriers could account for half the thromboembolic events seen in HRT users. Incidence rates for non-fatal cardiovascular disease within the Million Women Study are estimated to be about 0.3/100 women/year for ischaemic heart disease and 0.1/100 women/year for both stroke and venous thromboembolism. Follow-up of the entire cohort to the end of 2002 will yield an estimated 5400 non-fatal coronary events, 2700 non-fatal stokes and 3500 venous thromboembolic events.

Study Design

The susceptibility study will be a case-control study nested within the Million Women Study cohort. The study will include approximately 20 000 cases (women who have reported incident breast cancer (breast cancer diagnosed after recruitment to the Million Women Study) (10 000), or incident cardiovascular disease (or other diseases of interest)(10 000)) and 40 000 controls (women who have not reported the onset of breast cancer, cardiovascular disease or other diseases of interest since recruitment to the Million Women Study).

Biological samples

This study will involve the collection of biological samples for the analysis of DNA and for biochemical tests (such as measures of endogenous hormones). Blood samples would provide some advantages; plasma and/or serum would be available for biochemical analyses, and methods of DNA extraction and analysis and of sample storage are well established. If however it proves difficult to obtain blood samples from such a large number of women throughout the UK, self-administered buccal swabs provide a possible alternative. While DNA yields from buccal smears are generally lower than those from blood, and biochemical analyses on saliva samples may be more difficult than on plasma or serum, buccal smears have been used successfully for genetic testing in large epidemiological studies and methods for long-term storage are currently being tested with encouraging results. We will include both blood and buccal smear sampling in our pilot studies to ascertain which is most appropriate for the main study.

Eligibility

Women eligible for the susceptibility study will be those who have completed the primary follow-up questionnaire three years after recruitment (and in the case of women who have reported breast cancer, who have also completed a secondary breast cancer follow-up questionnaire). All women taking part in the MWS have given signed consent for follow-up.

Approach to participants
Eligible women will be sent a follow-up/susceptibility study questionnaire. The questionnaires will request further details of the woman's health and of disease risk factors such as use of hormone replacement therapy and family history of disease; questionnaires sent to case women will also ask for details of disease diagnosis and treatment. Two versions of the questionnaire will be used in the initial pilot study.

The version used to recruit women for blood collection will include a question asking if the woman might be willing in principle to supply a blood sample for the susceptibility study. The questionnaire will then be posted back to the Million Women Study co-ordinating centre. If the woman is willing, a letter and information sheet about the Million Women Study and the susceptibility study will then be sent to her general practitioner asking if they are willing to take the blood sample at their GP practice. If the GP does not object, the woman will be sent a letter and information sheet giving details of the susceptibility study, a blood collection kit with full instructions for the GP or Practice Nurse, and two copies of the consent form. She will be asked to attend the GP practice for blood collection and to post back the blood sample and signed consent forms to the Million Women Study co-ordinating centre in Oxford. The consent forms will be countersigned by a study investigator and one copy returned to the participant for her records.

With the other version of the questionnaire, used to recruit women for buccal smear collection, women will be sent a letter and information sheet explaining the susceptibility study and asking if they would consider providing a buccal cell/saliva sample for biochemical analyses and analysis of DNA, a kit for collecting the buccal swab, with full instructions, and two copies of the consent form. Women who decide to take part in the susceptibility study will be asked to take the buccal swab sample themselves at home and to post the sample and signed consent forms, with the follow-up questionnaire, to the Million Women Study co-ordinating centre in Oxford. The consent forms will be countersigned by a study investigator and one copy returned to the participant for her records.

Those who prefer not to provide a blood or buccal sample will be able to return the questionnaire alone, as part of continuing follow-up within the Million Women Study.

Processing and storage of data and of biological samples
Data from questionnaires will be entered onto a secure database at the Million Women Study co-ordinating centre. Initial processing and storage of blood and of buccal cell and saliva samples will take place at the Million Women Study co-ordinating centre laboratory. Further analyses, including the extraction of DNA and the analysis of genetic polymorphisms, will be carried out either at the Million Women Centre laboratory or at other secure specialist laboratories.

Validation of disease diagnoses
Diagnoses of breast cancer are validated within the existing structure of the Million Women Study by record linkage to participating National Health Service Breast Screening Programme centres, cancer registries and the Office of National Statistics. It will be necessary to validate self-reported diagnoses of cardiovascular disease by comparison with hospital records. This will initially be done for a sample of cases in a pilot study (see below); all case women taking part in the susceptibility study will be asked to give signed consent for us to contact their hospital consultant for further information if necessary.

Analysis
Standard statistical methods, including Mantel-Haenszel calculation of odds ratios and logistic regression modelling, will be used to obtain relative risks for the development of breast cancer, ischaemic heart disease, stroke and venous thromboembolism in relation to the separate and joint effects of HRT use and of relevant genetic polymorphisms. Analyses will take into account potential confounding factors such as previous medical history, tobacco and alcohol use; and such biochemical data as are available from the susceptibility study.

Power
Assuming a 60% response rate, we expect approximately 6000 cases of breast cancer and 1800 cases each of ischaemic heart disease, stroke and thromboembolism. With 3 controls per case, and

minimum population prevalences of 30% for HRT use and between 2 and 20% for relevant genetic polymorphisms, the study will have 95 % power at 99% significance to detect the following minimum relative risks: for breast cancer, for the effect of HRT exposure in the various genotype subgroups, 1.1-1.4; for the effect of each genotype in the various exposure subgroups, 1.2-1.4; and an interaction ratio (the extra effect of joint exposure to HRT and a genetic factor, above that expected under a multiplicative model) of 1.4; and for each cardiovascular disease, for effect of HRT exposure 1.3-6.0, for genotype 1.7-3.0, and an interaction ratio for joint exposure to HRT and Factor V Leiden of 2.0.

Consent, confidentiality and ethical issues

The study will conform to current ethical and legal guidelines regarding consent, confidentiality and the use of human biological samples, and accords with the protocol outlined for the UK Biobank study (www. ukbiobank.ac.uk/protocol.htm). It will be conducted in accordance with relevant aspects of the Data Protection Act, the Human Rights Act 1998, the General Medical Council's Guidance on Confidentiality and the Council of Europe's Recommendation on the Protection of Medical Data. It will follow the guidance outlined in the Medical Research Council's documents on Personal Information in Medical Research and on Collections of Human Tissue and Biological Samples for Use in Medical Research. Each biological sample will be physically identified only by a code number. The information needed to link this number with questionnaire data, which may include personal identifying data where necessary for future follow-up within the Million Women Study, will be held securely and separately from other databases. All data will be analysed only in anonymised form and future publications will not identify individual women taking part.

Participation in the susceptibility study will be entirely voluntary. Potential participants will be informed in writing that:

- their decision whether or not to participate will not affect their future health care in any way
- they are free to withdraw from the study at any time
- all information will be treated with absolute confidentiality in accordance with the Data Protection Act, used for medical research purposes only and will never be used in a way which could identify them personally
- they will not receive individual feedback about the results of genetic or other tests
- the study is important for future research and many of the tests and analyses which will be conducted in the future cannot be specified at present
- the study has been approved by the appropriate Multi- Research Ethics Committee.

Participants will be asked to give written consent for the storage of their blood or buccal cell/saliva sample and for the use of the sample for unspecified biochemical and genetic tests now and in the future.

Participants will be encouraged to ask about or comment on any aspect of the study either in writing or by telephoning a study Freephone number.

Pilot studies

Pilot studies will be carried out on a small sample (about 200 - 400 women with disease and a similar number of control women). The results, including comments from participants, will be used to test methods of collection, storage, processing and analysis of biological samples and to confirm acceptability and usefulness of the study methods for the questionnaires. A major feature of the pilot study will be to compare the feasibility of collection and analysis of blood samples and of buccal smears. As part of the pilot studies information from the hospital records of a sample of women reporting incident cardiovascular disease will be compared with self-reported information from questionnaires to assess diagnostic validity of questionnaire reports.

Publication of results

The Million Women Study is a joint research project of the NHS Breast Screening Programme, the Medical Research Council and Cancer Research UK. Results from the study will be published in peer-reviewed journals and appropriate authorities (eg the Department of Health and the Medicines Control Agency) will be notified of relevant results where necessary.

Summary of amendments to original protocol
The original protocol was approved in April 2004 (mws.dsw.revisedprotocol0310.doc) by Eastern MREC. To date the following amendments have been approved. Please note that, following pilot studies, we collect only blood samples for the main study, and not buccal (cheek) swabs.

1. Simplification of consent procedure (Approved 7.12.05)
It was agreed that we may send out 2 copies of the consent form with a printed signature from the Principal Investigator, and ask the participant to return one signed copy to us. This replaces the original procedure whereby we sent 2 unsigned copies, both were returned to us after signature by the participant and we then returned one, countersigned by hand, to the participant.

2. Change to recruitment procedure to allow use of up to date information on case/control status (Approved 20.1.06)
It was agreed that in order to minimise inappropriate mailing to participants, we will use information obtained from the current 2006 follow-up questionnaire for the main Million Women Study in order to identify potential participants as cases or controls for the Disease Susceptibility study. Approval was given for the 'recruitment' question asking if women might be willing to give blood to be moved from the Disease Susceptibility questionnaire (which will now be sent with the blood kit) to a separate insert sheet to be mailed with the general follow-up questionnaire.

3. Revisions to study documents to comply with above changes (Approved 10.3.06)
Minor revisions to approved documents were agreed, including consent form, participant information sheet, letters to GPs and questionnaire, in order to accommodate the changes in consent and recruitment; and changes to contact details due to our change of address in 2005.

4. Revision of GP letter and revision of Breast Cancer questionnaire (Approved 17.03.2006)
Introduction of letter explaining the study to all GPs in England to replace named GP letters. Revised copy of breast cancer questionnaire to bring in line with revised control questionnaire and remove questions asked in a previous questionnaire.

5. Revision of control questionnaire for England. Revision of documents for Scottish controls. (Approved 21.08.2006)
Revision of English control questionnaire to bring into line with breast cancer questionnaire. Revision of Scottish mailing procedure and documents to ask GP if willing to take blood sample for woman before sending blood kit to woman. Revision of Scottish documents to validate height and weight by adding height and weight question on questionnaire and on GP blood form. Adding height and weight request into individual Scottish GP letter.

6. Revision of cardiovascular disease questionnaire. (Approved 15.12.2006)
To bring questionnaire for women with cardiovascular disease in line with breast cancer questionnaire.

7. Minor document revisions and some document additions. Change of approach to GP's in England. (Approved 10.09.2007)
Following feedback from participants and GP practices we have:

- introduced letter to Practice Managers
- added insert to questionnaire explaining blood collection procedure
- changed some wording on cardiovascular case questionnaire
- updated GP letter
- updated participant information sheet
- updated blood sampling instructions sheet

Change of approach to GP's in England, to ask GP if willing to take blood sample for woman before sending blood kit to woman. Revision of English control and cardiovascular disease questionnaires and GP letter to accommodate this approach.

8. Revision of documents and questionnaire to validate self reported height, weight, waist,hip and blood pressure measurements in selected areas of England. (Approved 18.09.2008)

Revision of :

- Control questionnaire by adding height/weight/hip/waist/BP measurements questions.
- GP letter by asking to measure height/weight/hip/waist/BP, and offer re-imbursement.
- Practice Manager letter to explain new part of study as above, and reimbursement.
- Information sheet to add sentence to inform participants about extra measurements.
- Blood Form to inform practice re: reimbursement and redesign form with section to write results of measurements taken at the practice.

9. Resuming collection of blood samples without body measurements in England and Scotland (Approved 23.03.2009)

Revision of:

- Participant information Sheet to exclude measurement information
- GP letter to exclude measurement information and change re-imbursement amounts
- Practice Manager letter to explain invoicing procedure, change re-imbursement amounts, and contact details. Study progress updated.
- Blood sampling instructions to delete body measurement information and change amount for re-imbursement. Form structure change back to that previously used.

10. Validation of Self reported questionnaire data for vascular disease status by GP review of primary care medical data. (Approved 10.7.2009)

The following documents were approved, so that we can write to the GPs' of study participants asking for confirmation of self-reported presence and absence of vascular disease:

- Letter to GP explaining this part of the study and asking for their help
- Reply form 4 types:
    - IHD diagnosis
    - CVA diagnosis
    - VTE diagnosis
    - Vascular disease status

11. Addition of consultee letter to inform those participants' representatives who have given notice of a participant's mental incapacity of their options under the Mental Capacity Act. (Approved 18.2.2010)

The approval of these documents enable us to write to relatives or representatives of participants who have suffered a mental incapacity since the last time they were contacted for the study. The representatives will receive a letter outlining options, for continuing to use existing data we hold for the participant on the study, offering them the option of writing to Prof Beral to withdraw data already held for the participant. This complies with legislation under the Mental Capacity Act (2005)

12. Document and word revisions for Vascular Validation study. (Approved 04.03.2010)

- Minor wording revisions to letters based on GP feedback from pilot study
- Inclusion of a letter to Practice Managers to maximise the response rate
- Addition of coding boxes to reply forms
- Changes of wording on forms to enable clearer information to be collected.

13. Minor wording revisions for letters used in Vascular Validation study. (Approved 1.7.2010)

- Minor wording revisions:
    - Change of contact person at co-ordinating centre
    - Change of Payment arrangement details on Practice Manager letter

14 Letter to confirm funding arrangements until 2014 confirmed by MREC. (non substantial amendment, for information only) (Acknowleged 16.09.2010)

15 (REC amend 9) Change of study coordinator from Krys Baker to Lynden Guiver (non substantial amendment, for information only) (Acknowleged 01.05.2012)

16 Continuing collection of blood samples from women in the study. (Approved 03Dec2012 and revised 13Dec2012)

- Using the same method as before and with minor changes to some of the documents to bring them up to date

- BCa Q: MWS.CAB/200/1209, CTRL Q: CN1/200/0810, GP letter and information sheet, kit participation information sheet, practice manager letter, blood information leaflet, participant consent form, kit instructions.

17.Collection of blood samples from additional women who reported cardiovascular disease on any study questionnaire (Rather than just in the most recent questionnaire) (Approved 21.08.2015) – Minimal Dataset submission with IRAS

This extends the study of different cardiovascular conditions in order to study a rarer cardiovascular condition, aortic stenosis. The blood sampling collection method remains as approved in Amendment 9 (approved March 2009) and Amendment 16 (approved December 2012)

- Cardiovascular AS Questionnaire (MWS.CAC.200.1506.pdf)
- Leaflet to accompany questionnaire (DSW.info.leaflet.1209.pdf)
- GP letter (DSW.gp.letter.1506.doc)
- Practice Manager letter (DSW.pm.1506.doc)
- GP information sheet (DSW.gp.info.1506.doc)
- Patient letter and information sheet (DSW.kit.pt.info.1506.doc)
- Consent form (DSW.kit.consent.1209.doc)
- Blood sampling instructions for GP/Nurse (DSW.kit.inst.1506.doc)

NB: Admin change made to amend the GP info sheet,(GP.letter.info.1601) confirmed as non-substantial amendment (acknowleged 14.01.2016)

Non-substantial amendment (Acknowledged 28.10.2014) Letter confirming minor amendment- confirming continuation of funding from MRC until 30 Sept 2018

Non-substantial amendment (Acknowledged 29.10.2015): minor amendments to questionnaire and documentation for Aortic Stenosis.

Non-substantial amendment (Acknowledged 14.01.2016) to amend GP/consultation information sheets or letters for Aortic Stenosis.

Appendix 4: Million Women Study Advisory Committee: current membership and Terms of Reference (August 2017)

Million Women Study Advisory Committee

Membership:

Carol Dezateux (Chair)	University College London
Richard Peto	Clinical Trial Service Unit, Oxford
Valerie Beral	Million Women Study, co-Chief Investigator
Jane Green	Million Women Study, co-Chief Investigator
Julietta Patnick	University of Oxford
Emily Banks	Australian National University; 45 and Up cohort
Lucy Carpenter	Lay representative
Cathie Sudlow	University of Edinburgh; UK Biobank cohort

Terms of reference

1. To advise study investigators on scientific strategy for the Million Women Study.

2. To act as the Access Committee for requests for access to Million Women Study data, under the terms of the Million Women Study and Richard Doll Centenary Archive data access policies.

Appendix 5: Selected characteristics of participants and details of follow-up

Table 1. Characteristics of 1.32 million women at recruitment into the Million Women Study, 1996-2001

Year of birth, median (interquartile range)	1942 (1938-46)
Age at recruitment, years, mean (SD)	56 (5)
Nulliparous, %	11%
Number of children (in parous women), mean (SD)	2.4 (1)
Ever use of oral contraceptives, %	59%
Current smoker, %	20%
Does not drink alcohol, %	24%
Alcohol consumption (in drinkers), units/week (SD)	7.7 (8)
Height, mean cm (SD)	162 (7)
Weight, mean Kg (SD)	69 (13)
Body mass index, mean kg/m^2 (SD)	26.2 (5)
Any physical activity, 2 or more times per week, %	55%
Strenuous physical activity, more than once per week, %	21%
Menopausal hormones, % current user	33%

Table 2: Study re-survey questionnaires: timings and content

Self-reported postal questionnaire data						
		Recruitment	**3-year Re-survey**	**8-year Re-survey**	**12-year Re-survey**	**15-year Re-survey**
Year answered		1996-2001	1999-2005	2006-7	2010-12	2013-14
Data entry/checking		Complete	Complete	Complete	Complete	Ongoing
Number of women		1,320,000	869,000	688,000	604,000	Pending
Socio-demographic	Date of birth, country of birth, deprivation , education	•				
	Ethnicity, early life circumstances		•		•	•
Body size	Height	•		•	•	
	Weight (and derived BMI)	•	•	•	•	•
	Early life body size		•	•		
Lifestyle factors	Smoking	•	•	•	•	•
	Alcohol	•	•	•	•	•

	Physical activity	•	•	•	•	•
Past health		•	•	•	•	•
Family History		•	•	•	•	•
Reproductive factors	Menarche, age at each birth, breastfeeding, use of oral contraceptives	•				
	Menopause, hysterectomy	•	•	•		
	Use of menopausal hormones	•	•	•	•	•
Diet	Food groups, beverages		•	•	•	•
Wellbeing	Overall health, happiness, stress		•	•	•	•
Falls	Including fractures		•	•	•	•
Social & occupational	Social participation		•	•	•	•
	Work, shift work		•	•	•	•
	Marital status/ bereavement		•	•	•	•
	Sleeping, napping		•	•	•	•

Online questionnaires

Type	Date
24-hour dietary recall (Oxford WebQ)	2010 to present
24-hour recall of physical activity	2010 to present
Food frequency questionnaire	2012 to present
Use of e-cigarettes	2016 to present
Cognitive function: problem solving	2017
Changes to Lifestyle in Retirement (CLR)	2017

Table 3. Electronic linkages to routinely collected NHS and other health data

	Provider	**Population with data**	**Dates**	**Data**
Deaths	England: ONS/NHS Digital Scotland: ISDScotland	Whole cohort	Annual update	ICD-10 cause of death
Emigrations and other loss to follow-up	England: ONS/NHS Digital Scotland: ISDScotland	Whole cohort	Annual update	Reasons and dates
Cancer registrations	England: ONS/NHS Digital Scotland: ISDScotland	Whole cohort	Annual update	ICD-10 cancer site ICDO tumour morphology
Hospital inpatient/ day patient admissions	England: HES/NHS Digital Scotland: ISDScotland	Whole cohort	Annual update	ICD-10 diagnosis OPCS4 procedures
Cancer Screening	Public Health England	Participants in England	Up to 2013	Dates of invitations for breast, cervix and bowel screening and attendance
Cancer outcomes and services and dataset	Public Health England	Participants in England	Annual update	Additional data on tumour characteristics (e.g. stage, grade)
Primary care	England: Clinical Practice Research Datalink Scotland: ISD	100,000 women in England 100,000 women in Scotland	Up to 2013 2009-2015	Diagnosis and prescribing (Read/OXMIS codes) Dispensing of drugs prescribed in primary care

Table 4. Million Women Study: follow-up for deaths, cancer registrations and hospital admissions notified as of September 2016

Number of deaths (to 31 December.2015)	173,000
Number lost to follow-up (to 31 December 2015)	20,000
New cancer registrations (to 31 December 2014)	
Total number (excluding non-melanoma skin cancer)	198,000
Breast cancer	67,000
Colorectal cancer	18,000
Lung cancer	16,000
Non-melanoma skin cancer	51,000
Total number of hospital admissions (to 31 March 2015)	8,400,000
Women with at least one hospital admission:	1,150,000
for ischaemic heart disease	263,000
for stroke	84,000
for fracture	75,000

Appendix 6: Million Women Study Data Access and Sharing Policy

Nuffield Department of Population Health
Cancer Epidemiology Unit

Richard Doll Centenary Archive

Data Access and Sharing Policy for the Million Women Study

University of Oxford

V1.5 May 2018

Version History			
1.0	June 2015	Revised for consistency with Richard Doll Centenary Archive policy	Jane Green
1.1	September 2015	Updated for MRC report	Jane Green
1.2	September 2016	Updated in line with new NDPH RDCA policy	Jane Green
1.3	May 2017	Minor updates	Jane Green
1.4	August 2017	Revised advisory committee list	Jane Green
1.5	May 2018	Updated in line with revised NDPH RDCA policy	Jane Green

Contents

Introduction

This document describes the Data Access and Sharing Policy for the Million Women Study. Its purpose is to define policy and procedures for the study to allow appropriate data sharing for scientific research, and to ensure adherence to the common principles on data sharing outlined by Research Councils UK (1,2) and by the Expert Advisory Group on Data Access (3). Million Women Study data are held by the Cancer Epidemiology Unit (CEU), Nuffield Department of Population Health (NDPH), University of Oxford as part of the Richard Doll Centenary Archive. This policy is consistent with the Data Access and Sharing Policy of the Richard Doll Centenary Archive, and with NDPH Information Security Policies.

This policy does not cover Freedom of Information requests or data subject access requests under the Data Protection Act.

The aims of the Million Women Study are to study major modifiable causes of morbidity and mortality in women in middle and old age, by conducting high quality research that generates reliable results about the causes of, treatments for, and public health impact of, such diseases. The study welcomes data sharing that furthers these aims. Responsible sharing of publicly-funded research data is encouraged by major UK public and charitable funders, and means that the public health value of information such as that collected over many years by the Million Women Study can be maximised.

Core Principles for Data Access

The Million Women Study recognises the central principle that publicly-funded research data are a public good, produced in the public interest, and that such data should be appropriately preserved and made available for new research purposes in a timely and responsible manner.

Unregulated access to our study data is not possible. As was standard practice at the time, Million Women Study participants were not asked specifically for consent to data sharing with outside bodies at recruitment in 1996-2001. Information was received in confidence from study participants, and they were told that information and biological samples would be treated with absolute confidentiality and used only for medical research. We have the responsibility to ensure that data and samples are accessed only by bona-fide researchers of high scientific probity who have agreed to abide by the requirements described in this document and by any contractual arrangements with funders and with external suppliers of the data relevant to the dataset, and who have any necessary ethics and regulatory approvals in place.

Million Women Study policy on data access is based on the need to:

- **protect participants**, honour our commitments to them and act within the scope of their informed consent
- **ensure compliance** with legal and regulatory requirements (e.g. the Data Protection Act 1998 and the Human Tissue Act 2004)
- **prioritise access** to those parts of the resource which are limited in availability, especially the depletable resource of biological samples
- **foster high quality** research

Data and biological sample resources
This policy applies to all Million Women Study information resources held by the Cancer Epidemiology Unit (CEU), including the following categories:

- data on individual participants in epidemiological studies, including detailed personal information related to individuals' health; these data are derived partly from information received in confidence from participants, and partly from data received from NHS Central Registers, NHS Cancer Screening Centres, clinical records, and from other equivalent sources

- data, materials and results from analyses and investigations conducted using the resource biological samples

Million Women Study data are held securely and in accordance with the Data Protection Act and other legal and regulatory requirements. The Cancer Epidemiology Unit holds NHS IG Toolkit approval. Our data security and governance policies are available at: https://www.ceu.ox.ac.uk/policies2

Terms of access to Million Women Study data

As data custodian, the Million Women Study recognises its responsibility to maintain and preserve data for future use and to regulate data access. Data security is an integral part of our study, which conforms to the Cancer Epidemiology Unit Information Security and Governance protocols.

Access to data relating to individuals' health, lifestyle and environment, biological samples and data derived from samples will only be permitted by application and under a Data Transfer Agreement. The application process is summarised below (p6). Access will be permitted only for research use that is consistent with the purpose of the original research study, has as appropriate been ethically and scientifically approved by independent reviewers and is consistent with the existing study ethics framework and with participants' consent.

Identifying data will not be made available to external researchers. Identification of participants will be forbidden under the terms of the Data Transfer Agreement and we will take further steps to minimise risk of re-identification, supplying datasets labelled with study-specific pseudo-ID numbers, suppressing small numbers and providing data at the minimum level necessary for analyses (e.g. month/year of birth, not full date of birth).

The study investigators will not collaborate with any user or group of users on an exclusive basis, although access to the depletable parts of the resource will necessarily be limited. Prioritisation of requests for access to depletable biological samples will be determined according to criteria to be set by the study's Advisory Committee.

Applications for funding for research including data from the Million Women Study will be required to include a CEU study investigator as a co-applicant.

Privileged use of study data by Million Women Study researchers

The study team may claim exclusive use of data for their funded research during the period of funding (generally, 3-5 years) and for 12 months thereafter. Data are otherwise available for sharing, although the scale and complexity of the dataset and the time needed for data entry and validation/checking after data collection mean that there can be several years between acquiring data and their being ready to use in analyses. Data sharing through collaboration with the study team may allow external researchers to use study data sooner than this.

Ethics and Governance approvals

Obtaining ethical approval from the collaborator's ethics committee is the responsibility of the requester. The requester, in conjunction with the Million Women study investigators, may also need to obtain approval from the Research Ethics Committee responsible for the Million Women Study. Local Research Governance approvals, if required, are the responsibility of the requester. Approval(s) required from other bodies (e.g. NHS Digital or the Office for National Statistics, required for the sharing by the Million Women Study of linked health data are the responsibility of the requester, but applications should be made in conjunction with Million Women Study investigators.

Data transfer Agreement

Access to anonymised individual participant information or to biological samples is permitted only under a Data Transfer Agreement. This will specify the user and the specific purpose for which data will be made available. It will include standard terms as to ownership, exploitation and dissemination of results, including the return of results to the study for incorporation into the resource. It will specify the fee payable and include requirements that the user conform to Ethics and Governance terms for the secure management of confidential and sensitive information, the terms of the participants' consents and to this and to the Richard Doll Centenary Archive Data Access policies. Intellectual Property rights will be defined as appropriate for each Agreement.

Tabular data or specific requested analyses may be shared without a formal agreement if the Principal Investigators feel this is appropriate.

Fees

Collaborators are expected to contribute financially to the cost of data sharing (including preparation of tabular data); such contributions will include data access and service fees at a level no less than the administrative cost to the study of collaboration (preparation of data for sharing, etc.) and may by agreement be 'in kind'.

Publication of agreed data sharing projects

The Million Women Study, through the Richard Doll archive, will reserve the right to publish the title, the names(s) and affiliations(s) of the Chief Investigator(s), a lay summary and lay and scientific abstracts of each piece of collaborative research for which access to the resource has been granted, before identification or publication of results.

Notification of progress and feedback of enhanced data

Collaborators will be required to submit regular progress reports to the Million Women Study investigators. Accountability to funding bodies for collaborative projects is the responsibility of the grant holders i.e. the joint responsibility of the collaborator(s) and the Million Women Study as co-applicants.

All information (including results of analyses of biological samples) relating directly to study participants in all collaborative or external projects becomes part of the existing study. Users of study data will be required to provide the study with a copy of all the results from their research based on this material, including negative findings and supporting data, for incorporation into the resource under the same access terms (allowing for a period of protected access by the original collaborator) as those applying to the original data. Users of samples will be required to provide sufficient details of the assay techniques for other researchers to comprehend the results.

Publication of results

Where the Million Women Study is the sole or most substantial source of data for a publication, the study should be named in the Title, Subheading or Abstract of the paper.

Co-application for funding with a CEU research study investigator has no significance in itself in relation to named co-authorship of scientific papers. However in order to recognise the contribution made by past and current staff and collaborators (e.g. NHS Cancer Screening Centres) to setting up and maintaining study resources the Million Women Study team, and if appropriate individual investigators, will be named as co-authors on any publications.

Each paper to be submitted for publication by collaborators must be forwarded to the study investigators for consideration at least 28 days before submission.

Theses

We request that we are provided with an electronic copy of any thesis using Million Women Study data as soon as possible after a degree is awarded.

Media and press releases

All press releases on research arising from the study must be prepared in conjunction with the Million Women Study Investigators, or a Press team designated by them.

Constraints on data access and use

Data supplied will be used for the purposes of medical research only and within the constraints of the consent under which the data were originally gathered, and of any contractual agreements between the study from which data are requested and its funders or external data sources. Access will not be permitted for purposes other than scientific research (for example, for police or forensic use) except where required by court order.

Where demand for material exceeds its availability or staffing resources are insufficient to make data available, access will be prioritised by the Advisory and Oversight Committees on scientific merit.

Data or samples may not be transferred to third parties, i.e. those other than Requesters named at the time of the original application or in subsequent applications and specified in the Data Transfer Agreement or later amendments.

Application Process for access to Million Women Study data

Initial enquiries should be made by email through the Richard Doll Centenary Archive Data Access Coordinator richard.doll.archive@ceu.ox.ac.uk using the Million Women Study Preliminary Enquiry Form. A study Investigator will then contact requesters to discuss the application.

Requesters should be employees of a recognised academic institution, health service organisation or other research organisation with experience in medical research; and should be able to demonstrate, through their peer-reviewed publications in the area of interest, their ability to carry out the proposed study. Applications on behalf of students should be made by the research supervisor.

If the request proceeds, we will require details of the proposed study and of relevant funding on the full Data Access Application Form. Applications for any NHS linked data must be able to demonstrate the benefit of the use of these data for health and social care purposes. Applications may be sent for scientific review by independent peer reviewers. The request will be considered formally by the Million Women Study Principal Investigators and

by the study's Advisory Committee (Annex A). Their decisions are overseen by the Richard Doll Centenary Archive Data Access Oversight Committee (Annex B), which has independent scientific representation and which also acts as an appeals body for disputed decisions. As a study receiving Medical Research Council funding, the MRC is also able to advise in the case of disputed access.

An outline of the Data Access process is shown in Annex C.

References

1. Research Councils UK common principles on data policy: http://www.rcuk.ac.uk/research/datapolicy/
2. MRC Policy and Guidance on Sharing of Research Data from Population and Patient Studies http://www.mrc.ac.uk/Utilities/Documentrecord/index.htm?d=MRC008302
3. Expert Advisory Group on Data Access Report, Governance of Data Access. June 2015, Cancer Research UK, ESRC, MRC,Wellcome Trust: http://www.wellcome.ac.uk/About-us/Policy/Spotlight-issues/Data-sharing/EAGDA/wtp059350.htm

Annex A: Membership of the Million Women Study Advisory Committee

Chair	Professor Carol Dezateux, University College London
CEU members	Professor Dame Valerie Beral, Million Women Study Professor Jane Green, Million Women Study Professor Julietta Patnick, University of Oxford
External members	Professor Emily Banks, Australian National University; 45 & Up Study Dr Lucy Carpenter, Lay Member Professor Sir Richard Peto, University of Oxford Professor Cathie Sudlow, University of Edinburgh; UK Biobank
Secretariat	PA to Professor Valerie Beral

Annex B: Membership of the Richard Doll Centenary Archive Data Access Oversight Committee

Chair	Professor Dame Anne Mills (London School Hygiene & Tropical Medicine)
External Members	Dr Janet Valentine (Medicines and Healthcare products Regulation Agency) Professor Peter Sandercock (University of Edinburgh) Mr Jonathan Sellors (Legal Counsel, UK Biobank)
NDPH Members	Professor Jane Armitage (Clinical Trial Service Unit) Professor Sarah Parish (Clinical Trial Service Unit) Professor Jane Green (Cancer Epidemiology Unit) Professor Jenny Kurinczuk (National Perinatal Epidemiology Unit) Dr Michael Lay (NDPH Information Governance Lead)
Secretariat:	RDCA Administrator

Annex C: Flowchart of Data access process