

UNIVERSITY OF OXFORD

**The effects of childbearing on women's
body mass index, and on the risk of
diabetes mellitus, or ischaemic heart
disease after the menopause**

by

KLBobrow

A thesis submitted in partial fulfillment for the
degree of Doctor of Philosophy

in the
Department of Public Health
Medical Sciences Division

March 2012

Declaration of Authorship

I, Kirsten Bobrow, declare that this thesis titled, 'The effects of childbearing on women's body mass index, and on the risk of diabetes mellitus, or ischaemic heart disease after the menopause' and the work presented in it are my own. I confirm that:

- This work was done wholly or mainly while in candidature for a research degree at this University.
- Where any part of this thesis has previously been submitted for a degree or any other qualification at this University or any other institution, this has been clearly stated.
- Where I have consulted the published work of others, this is always clearly attributed.
- Where I have quoted from the work of others, the source is always given. With the exception of such quotations, this thesis is entirely my own work.
- I have acknowledged all main sources of help.
- Where the thesis is based on work done by myself jointly with others, I have made clear exactly what was done by others and what I have contributed myself.

Signed:

Date:

Abstract

Background: Excess adiposity, diabetes mellitus, and ischaemic heart disease are common important causes of morbidity and premature mortality in postmenopausal women in the UK. A large amount of data exists on known risk factors for these conditions, and for risk factors men and women share there is little evidence to suggest sex-based differences. It has been suggested that factors unique to women (such as parity and breastfeeding) may also influence risk. The nature of the relationship between childbearing and these conditions remains to be clarified. In this thesis I explore the association between women's childbearing histories and their adiposity, and risk of diabetes or ischaemic heart disease after the menopause, to provide evidence on the character, repeatability and public health relevance of the associations.

Aim: To explore the hypothesis that childbearing (specifically parity and breastfeeding) is associated with women's body weight and risk of excess adiposity, and also with women's risk of diabetes mellitus, and ischaemic heart disease after the menopause.

Methods: Data are analysed from a large population-based cohort of middle-aged UK women recruited in 1996 to 2001 (the Million Women Study) with complete childbearing information, and who had baseline anthropometry, and were followed for incident diabetes or ischaemic heart disease through repeat survey questionnaires, hospital admission records, and central registry databases.

Results: In a large ethnically homogeneous population of postmenopausal UK women increasing parity was associated with an increase in BMI, however this increase was offset in women who breastfed. The associations between parity, breastfeeding and BMI were of a similar order of magnitude to established risk factors known to be associated with BMI, for example smoking, and physical activity. The associations between childbearing and women's risk of diabetes mellitus after the menopause appear to be largely due to the effects of childbearing on maternal BMI. There is only limited evidence to suggest a direct effect of childbearing on women's risk of diabetes after the menopause. There is statistically significant evidence of an association between childbearing and women's risk of ischaemic heart disease after the menopause. Parity was associated with a modest increase in risk whereas breastfeeding was associated with a small decrease in risk, however the effects were small in comparison to known important risk factors.

Conclusions In a large population of UK women childbearing was found to have a persistent influence on women's mean BMI after the menopause, and through this postmenopausal risk of diabetes mellitus. Childbearing was also found to be modestly associated with women's risk of ischaemic heart disease after the menopause.

Acknowledgements

Although I am responsible for the contents of this thesis I was only able to carry out this large body of work with the explicit and implicit help of a welcoming research community and a public who believe in the value of such endeavors. In particular I would like to thank all of the UK and South African women who participated in either the Million Women or the Cardiovascular Risk in Black South Africans studies. I am extremely grateful to my supervisors Professor Dame Valerie Beral and Dr Maria Quigley for providing scientific advice and direction and for their patience, unfailing support and encouragement.

Funding for this DPhil project came from the Rhodes Trust and the Cancer Epidemiology Unit and I am indebted to the South African Rhodes selection committee, the Rhodes Trust, and Professor Dame Valerie Beral for ensuring I was always financially secure.

I thank my colleagues in the Cancer Epidemiology Unit and the Department of Public Health and Primary Care who have helped foster my development as a researcher by answering innumerable questions and providing insightful commentary to my work in progress.

In particular, I would like to thank Dr Gillian Reeves for statistical advice and reassurance, Dr Kirstin Pirie for facilitating access to the Million Women Study Data, Mrs Diana Bull for access to additional breastfeeding data, Ms Angela Balkwell for helping me understand the data and the on line study resources, double checking my numbers and code and for providing the data used in table 5.2 (page 147). I would also like to thank Angela Balkwell and Dr Delphine Casabonne for their assistance in Stata programming and for their good cheer and endless patience. I would also like to thank Mr Adrian Goodall for assistance in preparing the figures used in chapter 4 (page 87). I would like to thank Dr Lucy Wright and Dr Ben Cairns for access to and assistance in understanding the BMI validation study data. I thank Dr Miranda Armstrong for access to her results on the reliability of self-reported physical activity among Million Women Study participants. I am grateful to Mrs Barbra Crossley and Mrs Carol Keene for help in navigating the data and understanding the questionnaires. I would like to thank Dr Lucy Carpenter and Dr Vicky Benson for critically appraising my research and offering suggestions for improvements. I would like to thank Dr Bette Liu for

her practical support and access to findings from the 45andUp cohort. I would like to thank Dr Jo Watson and Dr Xian Sweetland for their advice and encouragement. I am indebted to Mr Dave Ewart and Mr Andrew Chadwick for helping me tame LaTeX, for rescuing accidentally deleted work, and reviving a dead laptop as well as for enduring the many other computer-related issues and ensuring all my computing requirements were well met. I would also like to thank Isobel, Joy, Pauline, Evie, Lindsay, Bryony, Kathy and the many other helpful individuals working in the Unit secretariat for their support and assistance.

I am most grateful to Dame Professor Valerie Beral, Dr Maria Quigley, Dr Gillian Reeves, and Dr Jane Green for reviewing the results of chapter 4 (page 87) and for practical assistance in preparing results for publication.

My research in South Africa would not have been possible without the assistance of Professor Naomi Levitt and Dr Krisela Steyn who allowed me to work on their study. I am extremely grateful to Sister Buyelwa Majikela-Dlangamandla and Sister Theresa Goga who translated and back-translated the questionnaire into isiXhosa, to Sister Serena Van Haght who worked with me to train interviewers, roll out the questionnaire and collect the data. I thank Sister Debbie Jonathan and Ms Erica April who very ably ran the clinics and ensured the reproductive histories questionnaire was completed. I also wish to thank the field-workers who recruited participants in a variety of demanding circumstances, Mrs Zukie Quesho, Mrs Rosalinda Tole, Ms Busisiwe Binase, Mrs Sabela Ndenliezi, Mr Zola Galela and Mr Diko. I am grateful to Mrs Sue Both for ably managing a very limited budget and for the innumerable tasks she performed which enabled the study to be completed.

Lastly I would like to thank my friends and family who have been incredibly patient during this process. I am extremely grateful to Dr Francesca Crowe for her encouragement and advice and for proof reading my thesis. I am also most grateful to my Room 25 office mate and partner in crime Dr Alison Price.

Contents

Declaration of Authorship	i
Abstract	ii
Acknowledgements	iii
List of Figures	ix
List of Tables	xii
Abbreviations	xv
1 Introduction	1
1.1 Rationale and aims	1
1.2 Outline of thesis	6
1.3 The role of the author in this thesis	8
2 The Million Women Study	18
2.1 Data sources and general methods	18
2.2 Sample	19
2.2.1 Recruitment	19
2.2.2 Resurvey	20
2.2.3 Follow-up	20
2.2.4 Ethics	21
2.3 Data definitions	21
2.3.1 Study Population	21
2.3.2 Exposures	22
2.3.3 Outcomes	25
2.3.4 Covariates	26
2.3.4.1 Defining Co-factors	27
2.3.4.2 Defining mediating Factors	29
Body mass index (BMI)	29
2.4 Data quality	30
2.4.1 Sample	30
2.4.1.1 Data entry & storage	30
2.4.1.2 Non-response & generalisability	31
2.4.2 Exposures	32
2.4.2.1 Parity	32

2.4.2.2	Breastfeeding	34
2.4.3	Outcomes	36
2.4.3.1	Body mass index (BMI)	37
2.4.3.2	Diabetes Mellitus	39
2.4.3.3	Ischaemic heart disease	39
2.4.4	Co-factors	41
2.4.4.1	Smoking	41
2.4.4.2	Physical activity	42
2.4.4.3	Alcohol	42
2.4.5	Exclusions and missing data	43
2.4.5.1	Exclusions	43
2.4.5.2	Missing data	44
2.5	Data analysis	47
2.6	Discussion	48
2.7	Chapter summary	49
3	Trends in breastfeeding in the United Kingdom	63
3.1	Introduction	65
3.2	Methods	65
3.2.1	Study Population	65
3.2.2	Data Definitions	67
3.2.3	Statistical Analysis	69
3.3	Results	70
3.4	Discussion	75
3.5	Chapter summary	78
4	Effects of childbearing on women's body mass index	87
4.1	Introduction	89
4.2	Methods	90
4.2.1	Study Population	90
4.2.2	Data Definitions	90
4.2.3	Statistical Analysis	92
4.3	Results	95
4.4	Discussion	101
4.5	Chapter summary	105
5	Effects of childbearing on women's risk of diabetes after the menopause	122
5.1	Introduction	124
5.2	Methods	125
5.2.1	Study Population	125
5.2.2	Data Definitions	126
5.2.3	Statistical Analysis	129
5.3	Results	132
5.4	Discussion	139
5.5	Chapter summary	144
6	Effects of childbearing on maternal risk of ischaemic heart disease after the menopause	168
6.1	Introduction	170
6.2	Methods	171
6.2.1	Study Population	171

6.2.2	Data Definitions	173
6.2.3	Statistical Analysis	175
6.3	Results	177
6.4	Discussion	184
6.5	Chapter summary	189
7	Childbearing and body mass index in a population of middle-aged (45 to 65 years) black South African women	208
7.1	Background and rationale	210
7.1.1	Associations between childbearing histories and BMI, and other risk factors for cardiovascular disease	210
7.1.2	Need for more evidence from different populations of women	210
7.1.3	Establishing research partnerships	211
7.2	Protocol for a sub-study within CRiBSA: Childbearing and body mass index a population of black South African women	212
7.2.1	The Cardiovascular Risk in Black South Africans study (CRiBSA)	212
7.2.2	Purpose of CRiBSA sub-study	212
7.2.3	Objectives	213
7.2.4	Design	213
7.2.5	Setting	213
7.2.6	Participants	213
7.2.7	Exposures	214
7.2.8	Main outcome measure	214
7.2.9	Methods	214
7.2.9.1	Data Analysis	214
7.2.9.2	Sample size	214
7.2.10	Questionnaire design, training materials, data collection and monitoring	216
7.2.10.1	Co-ordination and collaboration	216
7.2.10.2	Questionnaire design	216
7.2.10.3	Questionnaire pilot	217
7.2.10.4	Training materials	217
7.2.10.5	Data collection	217
7.2.10.6	Data monitoring and quality checks	218
7.2.10.7	Data capture and storage	218
7.3	Practical procedures	219
7.3.1	Recruitment	219
7.3.2	Questionnaire administration	220
7.4	Discussion	220
7.5	Chapter summary	220
8	Discussion	224
8.1	Summary of main findings	225
8.2	Possible implications	232
8.2.1	Younger women	233
8.2.2	Older women	233
8.3	Future research	234
8.3.1	Mechanism of action	234

8.3.2	Other outcomes	234
8.3.3	Generalisability of findings	235
8.4	Conclusions	235
A	Million Women Study questionnaires	241
A.1	Recruitment questionnaire (blue)	242
A.2	First repeat survey questionnaire (yellow)	246
B	Reliability of self-reported childbearing data in the Million Women Study	250
C	Cardiovascular Risk in black South Africans study (CRiBSA)	252
C.1	Summary study protocol: Cardiovascular Risk in black South Africans study (CRiBSA)	253
D	Questionnaires and training materials for childbearing patterns in and health profiles of a population of middle-aged black South African women	255
D.1	English-language version of questionnaire	256
D.2	isiXhosa-language version of questionnaire	260
D.3	English-language written training materials	264
E	Other work during this thesis	268
E.1	List of presentations during this thesis	268
E.2	Publications during this thesis	270

List of Figures

2.1	Location of NHS Breast Screening Programme centres before 1998	50
2.2	Histogram of parity among women in the Million Women Study	50
2.3	Distribution of total duration of breastfeeding among Million Women Study participants who breastfed	51
3.1	Histogram of calendar years in which parous Million Women Study participants reported having a birth	79
3.2	Trends in the proportion of UK women who breastfed over time using data from the Million Women Study, 1958 and 1970 birth cohorts, and the 1975, 1980, and 1985 National Infant Feeding surveys	83
3.3	Trends in the proportion of UK women who breastfed over time using data from the Million Women Study stratified by socioeconomic tertile as measured at recruitment into the study	83
3.4	Data from the Million Women Study showing the difference in the proportion of women who breastfed over time between women in the highest socioeconomic groups and the lowest socioeconomic groups	84
4.1	Flow diagram showing eligibility and inclusion criteria and the final number of women included in the main analyses	107
4.2	Scatter plot showing the association between BMIs calculated from measured height and weight data taken from a random sample of Million Women Study participants (x-axis) and self-reported height and weight data from the recruitment questionnaire (y-axis)	109
4.3	Bland-Altman plot showing relationship between the average BMI calculated from the two data sources (x-axis) and the difference between the BMIs calculated from the two data sources (y-axis)	109
4.4	Mean BMI (kg/m^2) among postmenopausal women by parity and breastfeeding standardised by age, region, socioeconomic group, smoking, and physical activity	112
4.5	Mean BMI (kg/m^2) among postmenopausal women by parity and breastfeeding standardised by age, region, socioeconomic group, smoking, and physical activity in upper, middle and lower socioeconomic tertiles . . .	113
4.6	Mean BMI (kg/m^2) among postmenopausal women by parity and increasing average duration of breastfeeding per child standardised by age, region, socioeconomic group, smoking, and physical activity	114
4.7	Change in standardised mean BMI (kg/m^2) among postmenopausal women for selected characteristics	115
5.1	Flow diagram showing eligibility and inclusion criteria and the final number of women included in the main analyses	145

5.2	Odds ratio of diabetes by parity in a model adjusted for age, region, and socioeconomic group (OR, 95%CI)	149
5.3	Odds ratio of diabetes among parous women by duration of breastfeeding in a model adjusted for age, region, and socioeconomic group (OR, 95%CI)	149
5.4	Odds ratio of diabetes by BMI in a model adjusted for age, region, and socioeconomic group (OR, 95%CI)	150
5.5	Odds ratio of diabetes by number of cigarettes smoked in a model adjusted for age, region, and socioeconomic group (OR, 95%CI)	150
5.6	Odds ratio of diabetes by category of strenuous physical activity in a model adjusted for age, region, and socioeconomic group (OR, 95%CI) .	151
5.7	Interaction effects of BMI (kg/m^2) on the odds ratios of diabetes associated with increasing parity and duration of breastfeeding,(OR, 95% CI) .	156
5.8	Odds ratio of diabetes by parity in the fully adjusted model (i.e. adjusted for age, region, smoking, BMI, alcohol, strenuous physical activity, and duration of breastfeeding as appropriate), (OR, 95% CI)	157
5.9	Odds ratio of diabetes among parous postmenopausal women by breastfeeding in the fully adjusted model, (OR, 95% CI)	158
5.10	Odds ratio of diabetes by BMI (kg/m^2) in the fully adjusted model, (OR, 95% CI)	159
5.11	Odds ratio of diabetes by smoking in the fully adjusted model, (OR, 95% CI)	159
5.12	Odds ratio of diabetes by socioeconomic group in the fully adjusted model, (OR, 95% CI)	160
5.13	Odds ratio of diabetes by strenuous physical activity in the fully adjusted model, (OR, 95% CI)	160
6.1	Flow diagram showing eligibility and inclusion criteria and the final number of women included in the main analyses	191
6.2	Incidence rate for IHD per 1000 women over a 5 year period by parity .	194
6.3	Incidence rate for IHD per 1000 women over a 5 year period among parous women by duration of breastfeeding	194
6.4	Incidence rate for IHD per 1000 women over a 5 year period by average number of cigarettes smoked per day	195
6.5	Incidence rate for IHD per 1000 women over a 5 year period by BMI . .	195
6.6	Incidence rate for IHD per 1000 women over a 5 year period by socioeconomic group	196
6.7	Incidence rate for IHD per 1000 women over a 5 year period by amount of strenuous physical activity per week	196
6.8	Incidence rate for ischaemic heart disease per 1000 women over a 5 year period among parous women by parity with adjustment for age, socioeconomic group, smoking, BMI, alcohol, strenuous physical activity, and duration of breastfeeding as appropriate, and region	199
6.9	Incidence rate for ischaemic heart disease per 1000 women over a 5 year period among parous women by duration of breastfeeding with adjustment for age, socioeconomic group, smoking, BMI, alcohol, strenuous physical activity, and parity as appropriate, and region	199

6.10	Incidence rate for ischaemic heart disease per 1000 women over a 5 year period among parous women by number of cigarettes smoked with adjustment for age, socioeconomic group, BMI, alcohol, strenuous physical activity, parity, and duration of breastfeeding as appropriate, and region	200
6.11	Incidence rate for ischaemic heart disease per 1000 women over a 5 year period among parous women by BMI with adjustment for age, socioeconomic group, smoking, alcohol, strenuous physical activity, parity, and duration of breastfeeding as appropriate, and region	200
6.12	Incidence rate for ischaemic heart disease per 1000 women over a 5 year period among parous women by socioeconomic group with adjustment for age, smoking, BMI, alcohol, strenuous physical activity, parity, and duration of breastfeeding as appropriate, and region	201
6.13	Incidence rate for ischaemic heart disease per 1000 women over a 5 year period among parous women by amount of strenuous physical activity per week with adjustment for age, socioeconomic group, smoking, BMI, alcohol, parity, and duration of breastfeeding as appropriate, and region	201

List of Tables

2.1	Summary of published studies looking at reliability of long-term maternal recall of parity	52
2.2	Summary of published studies looking at reliability of long-term maternal recall of parity cont.	53
2.3	Summary of published studies looking at reliability of long-term maternal recall of parity cont.	54
2.4	Summary of published findings on the validity of long-term maternal recall of breastfeeding	55
2.5	Summary of published findings on the reliability of long-term maternal recall of breastfeeding	56
2.6	Self-reported diabetes among Million Women Participants as reported at recruitment and at first repeat survey (about 3 years after recruitment) .	57
3.1	Breastfeeding questions from the Million Women Study recruitment questionnaire (aqua), the 1958 and 1970 birth cohort questionnaires, and the 1975, 1980, and 1985 National Infant Feeding survey questionnaires . . .	80
3.2	Characteristics of studies used in this analysis - 1958 and 1970 Birth Cohorts and the 1975, 1980, and 1985 National Infant Feeding surveys . . .	81
3.3	Characteristics of studies used in this analysis - the Million Women Study	81
3.4	Comparison of estimates of the proportion of women breastfeeding, and of the estimates by birth order for comparable years between the Million Women Study and the 1958, and 1970 birth cohorts, and the 1975, 1980, and 1985 National Infant Feeding surveys	82
3.5	For women who breastfed, comparison of categories of breastfeeding duration for comparable years between the Million Women Study and the 1958, and 1970 birth cohorts	82
4.1	Characteristics of study participants used in the analysis of body mass index by parity and breastfeeding status	108
4.2	Association between parity and breastfeeding and mean BMI (kg/m^2) among postmenopausal women in a model adjusted only for age and region	110
4.3	Change in mean BMI (kg/m^2) among postmenopausal women among parous women by parity and total duration of breastfeeding in models variously standardised	111
4.4	Change in standardised mean BMI (kg/m^2) among postmenopausal women by parity and total duration of breastfeeding for a randomly selected sample of 2,800 women with BMIs calculated from measured heights and weights	116

4.5	Change in standardised mean BMI (kg/m^2) among postmenopausal women by parity and total duration of breastfeeding using BMI data from women with measured heights and weights and applying it to the entire study population using the regression dilution approach [?]	117
5.1	Characteristics of the study population according to parity and breast-feeding status	146
5.2	Selected recruitment characteristics of women who did and who did not return a completed first repeat survey questionnaire	147
5.3	Odds ratio of diabetes among postmenopausal women for selected characteristics in models adjusted only for age and region (OR, 95%CI)	148
5.4	Odds ratio of diabetes among postmenopausal women associated with parity and breastfeeding in models adjusted for various important co-factors (OR, 95%CI)	152
5.5	χ^2 statistic for heterogeneity across categories of parity (three degrees on freedom) and durations of breastfeeding (three degrees of freedom) for each of the models in table 5.4 (page 152)	153
5.6	Odds ratio of diabetes among parous postmenopausal women associated with parity and breastfeeding in a model adjusted for age, region, smoking, socioeconomic group, BMI, alcohol, and strenuous physical activity (OR, 95% CI)	154
5.7	Odds ratio of diabetes among parous postmenopausal women associated with increasing parity and duration of breastfeeding stratified by category of BMI (kg/m^2), (OR, 95% CI)	155
5.8	Odds ratio of diabetes among postmenopausal women associated with parity and duration of breastfeeding in the fully adjusted model and in models which have additional adjustments made for ethnicity, family history of diabetes, and mother's own birth-weight (kg), (OR, 95% CI) .	161
5.9	Odds ratio of diabetes among parous postmenopausal women associated with parity and duration of breastfeeding in the fully adjusted model (age, region, socioeconomic group, BMI, smoking, alcohol, strenuous physical activity, and parity and breastfeeding as appropriate) and in a model which is similarly adjusted but where the category mean for BMI was calculated using measured heights and weights from a random sample of women (taken about 8 years after recruitment), (OR, 95% CI)	162
6.1	Characteristics of study participants included in the main analyses of the effects of childbearing on women's risk of ischaemic heart disease according to parity and breastfeeding status	192
6.2	Risk of an IHD event for selected characteristics in models adjusted only for age and region (HR, 95%CI)	193
6.3	Relative risks of ischaemic heart disease among postmenopausal women associated with parity and breastfeeding in models adjusted for various important co-factors (RR,95%CI)	197
6.4	χ^2 statistic for heterogeneity across categories of parity (three degrees on freedom) and durations of breastfeeding (three degrees of freedom) for each of the models in table 6.3 (page 197)	198
6.5	Risk of ischaemic heart disease among parous women in a model adjusted for age, socioeconomic group, smoking, BMI, alcohol, strenuous physical activity, and region (HR,95%CI)	202

6.6	The effect on model estimates of including women who had had a hospital admission for IHD prior to study recruitment, of only including women with complete information on all relevant factors, and of excluding the initial 3 years of follow-up (HR, 95%CI)	203
6.7	The effect on model estimates of adding additional selected co-factors (HR, 95%CI)	204
7.1	Sample sizes required to adequately power a study designed to detect a change in the mean BMI between two exposure groups	222
7.2	Selected characteristics of women participating in the control arm of a Cape Town-based case-control study of the association between lactation and risk of breast cancer (Coogan et al 1999)	222

Abbreviations

BMI	Body Mass Index
CI	95% Confidence Interval
CRiBSA	Cardiovascular Risk factors in Black South Africans
DM	Diabetes Mellitus
HES	Hospital Episode Statistics
HR	Hazard Ratio
ICD10	International Classification of Diseases 10th revision
IHD	Ischaemic Heart Disease
NHS	National Health Service
OR	Odds Ratio
RR	Relative Risk
UK	United Kingdom
WHO	World Health Organisation

Chapter 1

Introduction

1.1 Rationale and aims

Childbearing and women's long-term health and risk of disease

Childbearing (specifically parity and breastfeeding) is known to affect women's health in the short and longer term. [1–4] Among women whose reproductive careers are over investigators have found that childbearing affects women's longevity, as well as their risks of certain diseases including common causes of premature death. [1, 5–8] For example investigators have shown persistent protective associations between childbearing and women's risk of cancer of the breast and other reproductive organs i.e. endometrium and ovary, [1, 9, 10] Investigators have also found that childbearing is associated with risk of symptomatic gallbladder disease and cholecystectomy (risk increases the more children a woman has had, but decreases the longer she breastfeeds). [11] Much less is known about the effects of childbearing on risk for other common conditions which may also be influenced by reproductive hormones. Excess adiposity (i.e. overweight and obesity as captured by body mass index), diabetes mellitus (type

II), and ischaemic heart disease are three such conditions that are increasingly common, are associated with excessive morbidity and premature mortality, and are also associated with rising health care costs in the UK. [12–15]

In the UK, as in other developed countries, the population of older women at risk of these conditions is growing rapidly. [16] Older age, smoking, hypertension, dyslipidaemia, sedentary lifestyle and poor diet are known to be important risk factors for these conditions in this population. [17–30] Although results from epidemiological studies and hormone replacement therapy trials suggest that the risks of these conditions may be influenced by reproductive hormones, [31] a comprehensive review of the published literature using the terms “childbearing”, “maternal”, “postpartum”, “parity”, “pregnancy”, “breastfeeding” separately and in combination, and broad terms for each of the outcomes of interest (excess adiposity, diabetes, and cardiovascular disease) revealed that the extent to which childbearing is associated with women’s risk of excess adiposity, diabetes, and cardiovascular diseases is yet to be fully understood.

Body weight and excess adiposity

Overall adiposity is reasonably well captured by the measure body mass index (BMI; weight (kg)/height (m²)). Excess adiposity conventionally refers to BMIs ≥ 25 kg/m² with a BMI between 25 kg/m² to 29.9 kg/m² defined as “overweight” and BMI ≥ 30 kg/m² as “obese” (with various sub-groups defined for BMIs above this level.) [32] Raised BMI is a known and important risk factor for diabetes mellitus (type II), and heart disease, and to a lesser extent cancer. [12, 17, 30, 33–35] It has also been shown that excess adiposity is an independent risk factor for premature death. [29]

The relationships between childbearing and maternal adiposity are complex and unclear. Investigators have found that among reproductive-aged women in developed countries weight, in the short-term, tends to increase with each birth. [36–38] The published evidence of the short-term effects of breastfeeding on body weight is less consistent. [4, 39] In the longer term, after a woman's reproductive years are over, some but not all investigators have reported that women's adiposity (variously measured) increases with the number of children they have had. [40–45] [32] To date, there is very little published evidence on the relationship between women's breastfeeding histories and subsequent adiposity in later years and the findings are unclear. [46, 47] There is also little published evidence on the relative effect size of childbearing on women's adiposity in later years as compared to known important co-factors.

Diabetes mellitus (type II)

Diabetes mellitus (type II) is a common chronic condition among older UK women responsible for a substantial proportion of ill health and premature death. [14, 48–50] Well established risk factors for diabetes mellitus (type II) among older women include having a higher BMI, being sedentary, being a current smoker, and having a poor diet. [14, 32, 34, 50] It has been suggested childbearing may influence women's risk of diabetes mellitus (type II) after her reproductive years are over however the existing evidence is sparse and unclear. [51–58] There is also little information on the size of the effects of childbearing on women's risk of diabetes mellitus (type II) relative to known important risk factors such as BMI.

Ischaemic Heart Disease

Ischaemic heart disease (IHD) is another major cause of premature death in postmenopausal women in the UK and elsewhere. [59] Established risk factors for IHD include smoking, excess adiposity, high blood pressure, high blood cholesterol, and sedentary lifestyle. [17, 18, 20, 60, 61] Data from mortality surveys and cross-sectional analyses suggest that increasing parity is associated with an increased likelihood of ischaemic heart disease. [1, 5, 62–64] The evidence for an effect of breastfeeding on women's risk of ischaemic heart disease is equivocal. [65, 66] There are no published data on the independent (and joint) effects of parity and breastfeeding on women's risk of ischaemic heart disease, and their relevance as compared to established risk factors.

Question addressed in this thesis

Aim

The primary aim of this thesis is to explore the hypothesis that childbearing (specifically parity and breastfeeding) is associated with women's body weight and risk of excess adiposity, and also with women's risk of diabetes mellitus (type II), and ischaemic heart disease after the menopause.

Objectives

- Review and assess the quality and appropriateness of the data to be analysed
- Explore the associations between childbearing and BMI as excess adiposity is a known risk factor for both diabetes mellitus (type II), and ischaemic heart disease

- Explore the associations between childbearing and risk of diabetes independent of socioeconomic group, excess adiposity, smoking, and other important co-factors
- Explore the associations between childbearing and ischaemic heart disease independent of socioeconomic group, excess adiposity, smoking, and other important co-factors
- Compare the size of the effects of childbearing on these outcomes relative to the effect size of known important co-factors
- Design a small study to explore the nature and reproducibility of these findings in a different population of women

Thesis plan

Other than the introduction, this thesis consists of one chapter describing the data and methods used, followed by one chapter exploring breastfeeding trends in the UK, three chapters of analysis using data from the Million Women Study, and one chapter introducing and describing the protocol for a study of women's reproductive histories in a developing country context. Thereafter a final chapter will summarise work done, discuss major findings, and highlight future research opportunities. Each of the chapters broadly follow the same format; general introduction, discussion of data and methods of analysis used, presentation of results, followed by a discussion. Brief chapter summaries are presented below.

1.2 Outline of thesis

Chapter 1 provides a general introduction to the work presented in this thesis. The rationale for this research project is described and the overall aim and specific objectives of the thesis are introduced. The structure of the thesis and brief chapter outlines are presented. A detailed discussion of the role of the authors is also presented.

Chapter 2 introduces the data sources and general statistical methods used in this thesis. As the bulk of work presented relies on data from the Million Women Study a detailed exploration and discussion of the quality of this data are presented.

Chapter 3 describes historical trends in the annual proportion of women who breast-fed in the UK during the second half of the 20th century using data from the Million Women Study, and compares these to estimates from two birth cohorts and three infant feeding surveys. This chapter builds on the detailed discussion of the available breastfeeding data presented in chapter 2 (page 18) demonstrating how this data can be used to assess trends in breastfeeding over time and comparing it with other available sources of breastfeeding data.

Chapter 4 explores the association between childbearing and women's body mass index (BMI) after the menopause taking into account socioeconomic group and other important co-factors. Findings of the relative effects of parity and of breastfeeding on women's self-reported BMI are presented. In addition, the use of self-reported data to calculate BMI is discussed in detail and results are presented of sensitivity analyses exploring how misclassification of BMI based on self-reported height and weight data may affect these results.

Chapter 5 explores the association between childbearing and women's risk of diabetes mellitus after the menopause taking into account BMI and other important co-factors. Again building on findings from chapters 2 and 4 (pages 18, 87) the independent and joint associations between parity and breastfeeding and women's risk of diabetes are described, as are the associations between known important co-factors and women's risk of diabetes. Based on the findings in chapter 4 (page 87) the role of BMI as an effect mediator is explored. The relative size of the effects on women's risk of diabetes associated with childbearing and other important co-factors are presented and discussed.

Chapter 6 explores the association between childbearing and women's risk of incident heart disease after the menopause taking into account known important risk factors such as smoking. Similar in format to chapter 5 (page 122) the independent and joint associations between parity and breastfeeding, and also between important known co-factors and women's risk of heart disease are described. Again the role of body mass index in the potential associations between childbearing and risk of heart disease is explored. The relative size of the effects on women's risk of heart disease associated with childbearing and other important co-factors are presented and discussed.

Chapter 7 introduces a study protocol to explore the association between childbearing and BMI in a population of middle-aged black South African women. The chapter begins by considering the background and rationale for such a study. The design and development of an appropriate questionnaire are discussed. Also discussed are issues around sample size calculations and practical procedures including recruitment, questionnaire administration, quality checks, data capture and data storage.

Chapter 8 reviews the main findings from the thesis and discusses their public health relevance. The chapter begins with a brief review of the central findings from each chapter. For the results chapters there is also a comparative analysis of findings and further interpretation of the public health relevance of the results. Finally recommendations for further work are considered.

1.3 The role of the author in this thesis

I carried out all of the work presented in this thesis under the supervision of Dr Maria Quigley and Professor Dame Valerie Beral. The initial idea of exploring the associations between parity and breastfeeding and women's health after the menopause was first suggested to me by my supervisor Professor Dame Valerie Beral. In reviewing the published literature on the associations between parity and breastfeeding and women's health after the menopause I realised that despite a growing literature suggesting that childbearing influenced women's risk of chronic disease after the menopause the existing evidence was based on relatively small studies and the results were conflicting. In particular, the evidence for the long-term association between women's breastfeeding histories and their risk of heart disease and related diseases was limited. In consultation with my supervisors I decided to explore the associations between women's parity and breastfeeding and their risks of cardiovascular and related diseases within the context of the Million Women Study.

As the data on childbearing and breastfeeding exposures in the Million Women Study is collected via self-report and dependent on long-term recall I began by reviewing the known literature on the reliability and validity of such data. I obtained access to birth cohort data through the UK Data Archive and infant feeding surveys and used these to

compare how information on breastfeeding available from the Million Women Study compared to national breastfeeding rates in specific years. I also reviewed the published evidence on the reliability and validity of the sources of outcomes data I planned to use, and where necessary explored supplementary sources of data (for example, I made use of the GP-measured height and weight data kindly provided by Dr Lucy Wright and Dr Ben Cairns).

Having established the utility of the exposure and outcome data I performed analyses on the association between parity and breastfeeding and women's adiposity (as captured by BMI) after the menopause. Based on these findings and the importance of excess adiposity as a risk factor for cardiovascular and related diseases I decided to explore the associations between parity and breastfeeding and women's risk of heart disease and type II diabetes. In particular I wanted to explore whether an association between parity and breastfeeding and these outcomes existed and, if so, whether the association was independent of adiposity.

As an extension of the research and analyses I carried out using data from the Million Women Study I wanted to explore whether similar associations would be found in different populations of women, in particular women from less economically developed countries. As there are no available data from developing-economy countries I explored the possibility of undertaking a small cross-sectional study of childbearing and women's health profiles in middle-age in South Africa.

In 2008 I established contact with researchers at the University of Cape Town and the Africa Centre to explore the possibility of establishing a study of the health of post-menopausal women in relation to their reproductive histories. I visited both sites and discussed my study proposal with senior scientists.

The Africa Centre is a demographic and health surveillance unit in northern Kwa-Zulu Natal funded through the Wellcome Trust. The Centre is focused on HIV-related research. Data on reproductive aged women including prospectively collected breastfeeding information was available from the Africa Centre. Unfortunately, it was beyond the scope of this thesis to explore the available data on the short term effects of parity and breastfeeding on women's adiposity.

The Department of Medicine at the University of Cape Town (UCT) has an established research record, and has published data on cardiovascular risk factors in men and women living in urban settings. [67, 68] In 2008 Professor Naomi Levitt and Dr Krisela Steyn from the Department of Medicine at University of Cape Town and in collaboration with the Medical Research Council of South Africa established a population-based survey to update prevalence data on cardiovascular disease and associated risk factors. [69] Within the given resource and time constraints we established that it would be feasible to add a questionnaire on women's reproductive histories to the existing study. In consultation with senior scientists in the Unit and at the University of Cape Town I developed and designed an interviewer-administered questionnaire and protocol for a small cross-sectional study within the larger population-based survey. With the assistance of Sister Buyelwa Majikela-Dlangamandla and Sister Theresa Goga who translated the questionnaire and provided valuable feedback I piloted and refined the questionnaire. I also developed appropriate training materials for interviewers.

Using data from the Million Women Study (MWS), the birth cohort and infant feeding surveys, and the Cardiovascular Risk in Black South Africans study (CRiBSA) I conducted all of the statistical analyses presented in this thesis. I conceived and developed the ideas for each of the chapters and I wrote all of the chapters in this thesis including

the scientific paper submitted for publication and the conference presentation. My co-authors for the published abstract and submitted paper (Professor Dame Valerie Beral, Dr Jane Green, Dr Maria Quigley and Dr Gillian Reeves) read and provided advice on the statistical methods and content of each. The submitted paper has also been edited to take into account comments from the peer review process.

Bibliography

- [1] Beral V. Long term effects of childbearing on health. *J Epidemiol Community Health*. 1985 Dec;39(4):343–346.
- [2] Dugdale AE, Eaton-Evans J. The effect of lactation and other factors on post-partum changes in body-weight and triceps skinfold thickness. *British Journal of Nutrition*. 1989 Mar;61(2):149–153.
- [3] Glazener CMA, Abdalla M, Stroud P, Templeton A, Russell IT, Naji S. Postnatal maternal morbidity: extent, causes, prevention and treatment. *BJOG: An International Journal of Obstetrics and Gynaecology*. 1995;102(4):282–287.
- [4] Ohlin A, Rossner S. Factors related to body weight changes during and after pregnancy: the Stockholm Pregnancy and Weight Development Study. *Obesity research*. 1996 May;4(3):271–276.
- [5] Green A, Beral V, Moser K. Mortality in women in relation to their childbearing history. *BMJ: British Medical Journal*. 1988 Aug;297(6645):391–395.
- [6] Doblhammer G. Reproductive history and mortality later in life: A comparative study of England and Wales and Austria. *Population Studies: A Journal of Demography*. 2000 Jul;54(2):169–176.
- [7] Doblhammer G, Oeppen J. Reproduction and longevity among the British peerage: the effect of frailty and health selection. *Proceedings of the Royal Society B: Biological Sciences*. 2003 Aug;270(1524):1541–1547.
- [8] Grundy E. Women's fertility and mortality in late mid life: A comparison of three contemporary populations. *American Journal of Human Biology: The Official Journal of the Human Biology Council*. 2009 Aug;21(4):541–547.
- [9] Adami H, Lambe M, Persson I, Ekbom A, Adami HO, Hsieh CC, et al. Parity, age at first childbirth, and risk of ovarian cancer. *The Lancet*. 1994 Nov;344(8932):1250–1254.
- [10] Collaborative Group on Hormonal Factors in Breast Cancer. Breast cancer and breast-feeding: collaborative reanalysis of individual data from 47 epidemiological studies in 30 countries, including 50 302 women with breast cancer and 96 973 women without the disease*. *The Lancet*. 2002 Jul;360(9328):187–195.
- [11] Liu B, Beral V, Balkwill A. Childbearing, breastfeeding, other reproductive factors and the subsequent risk of hospitalization for gallbladder disease. *International Journal of Epidemiology*. 2009 Feb;38(1):312–318.

- [12] Murray CJ, Lopez AD. Global mortality, disability, and the contribution of risk factors: Global Burden of Disease Study. *The Lancet*. 1997 May;349(9063):1436–1442.
- [13] Allender S, Rayner M. The burden of overweight and obesity-related ill health in the UK. *Obesity Reviews*. 2007;8(5):467–473.
- [14] Chaturvedi N. The burden of diabetes and its complications: Trends and implications for intervention. *Diabetes Research and Clinical Practice*. 2007 May;76(3, Supplement 1):S3–S12.
- [15] O’Flaherty M, Ford E, Allender S, Scarborough P, Capewell S. Coronary heart disease trends in England and Wales from 1984 to 2004: concealed levelling of mortality rates among young adults. *Heart*. 2008 Feb;94(2):178–181.
- [16] Murray CJ, Lopez AD. Alternative projections of mortality and disability by cause 1990–2020: Global Burden of Disease Study. *The Lancet*. 1997 May;349(9064):1498–1504.
- [17] Yusuf S, Hawken S, Ounpuu S, Dans T, Avezum A, Lanas F, et al. Effect of potentially modifiable risk factors associated with myocardial infarction in 52 countries (the INTERHEART study): case-control study. *Lancet*. 2004 Sep;364(9438):937–952.
- [18] Doll R, Peto R, Boreham J, Sutherland I. Mortality in relation to smoking: 50 years’ observations on male British doctors. *BMJ*. 2004 Jun;328(7455):1519.
- [19] Bernstein MS, Costanza MC, Morabia A. Association of physical activity intensity levels with overweight and obesity in a population-based sample of adults. *Preventive Medicine*. 2004 Jan;38(1):94–104.
- [20] Lewington S, Clarke R. Combined Effects of Systolic Blood Pressure and Total Cholesterol on Cardiovascular Disease Risk. *Circulation*. 2005 Nov;112(22):3373–3374.
- [21] Rimm EB, Chan J, Stampfer MJ, Colditz GA, Willett WC. Prospective study of cigarette smoking, alcohol use, and the risk of diabetes in men. *BMJ*. 1995 Mar;310(6979):555–559.
- [22] Mucha L, Stephenson J, Morandi N, Dirani R. Meta-analysis of disease risk associated with smoking, by gender and intensity of smoking. *Gender Medicine*. 2006 Dec;3(4):279–291.
- [23] Smith J. Multiple Risk Factors for Cardiovascular Disease and Diabetes Mellitus. *The American Journal of Medicine*. 2007 Mar;120(3, Supplement 1):S3–S11.
- [24] LaRosa JC. Triglycerides and Coronary Risk in Women and the Elderly. *Arch Intern Med*. 1997 May;157(9):961–968.
- [25] Anand SS, Islam S, Rosengren A, Franzosi MG, Steyn K, Yusufali AH, et al. Risk factors for myocardial infarction in women and men: insights from the INTERHEART study. *European Heart Journal*. 2008 Apr;29(7):932–940.
- [26] Hart CL, Hole DJ, Lawlor DA, Smith GD. How many cases of Type 2 diabetes mellitus are due to being overweight in middle age? Evidence from the Midspan prospective cohort studies using mention of diabetes mellitus on hospital discharge or death records. *Diabetic Medicine: A Journal of the British Diabetic Association*. 2007 Jan;24(1):73–80.

- [27] Spencer EA, Pirie KL, Stevens RJ, Beral V, Brown A, Liu B, et al. Diabetes and modifiable risk factors for cardiovascular disease: the prospective Million Women Study. *European Journal of Epidemiology*. 2008;23(12):793–799.
- [28] Weinstein AR, Sesso HD, Lee I, Rexrode KM, Cook NR, Manson JE, et al. The Joint Effects of Physical Activity and Body Mass Index on Coronary Heart Disease Risk in Women. *Arch Intern Med*. 2008;168(8):884–890.
- [29] Whitlock G, Lewington S, Sherliker P, Clarke R, Emberson J, Halsey J, et al. Body-mass index and cause-specific mortality in 900 000 adults: collaborative analyses of 57 prospective studies. *Lancet*. 2009;373(9669):1083–1096.
- [30] Biggs ML, Mukamal KJ, Luchsinger JA, Ix JH, Carnethon MR, Newman AB, et al. Association Between Adiposity in Midlife and Older Age and Risk of Diabetes in Older Adults. *JAMA*. 2010 Jun;303(24):2504–2512.
- [31] Beral V, Banks E, Reeves G. Evidence from randomised trials on the long-term effects of hormone replacement therapy. *Lancet*. 2002 Sep;360(9337):942–944.
- [32] World Health Organisation W. Global Database on Body Mass Index [online]; 2010. Internet [cited 02 November 2010]. Available from: http://apps.who.int/bmi/index.jsp?introPage=intro_3.html.
- [33] Fuller JH, Stevens LK, Wang S. Risk factors for cardiovascular mortality and morbidity: The WHO multinational study of vascular disease in diabetes. *Diabetologia*. 2001;44(14):S54–S64.
- [34] Hu FB, Manson JE, Stampfer MJ, Colditz G, Liu S, Solomon CG, et al. Diet, Lifestyle, and the Risk of Type 2 Diabetes Mellitus in Women. *New England Journal of Medicine*. 2001;345(11):790–797.
- [35] Taylor AE, Ebrahim S, Ben-Shlomo Y, Martin RM, Whincup PH, Yarnell JW, et al. Comparison of the associations of body mass index and measures of central adiposity and fat mass with coronary heart disease, diabetes, and all-cause mortality: a study using data from 4 UK cohorts. *The American Journal of Clinical Nutrition*. 2010;.
- [36] Sheldon JH. Maternal obesity. *Lancet*. 1949 Nov;2(6585):869–873.
- [37] Harris HE, Ellison GT, Holliday M. Is there an independent association between parity and maternal weight gain? *Annals of Human Biology*. 1997 Dec;24(6):507–519.
- [38] Gunderson EP, Abrams B. Epidemiology of Gestational Weight Gain and Body Weight Changes After Pregnancy. *Epidemiologic Reviews*. 1999 Jan;21(2):261–275.
- [39] Sichieri R, Field AE, Rich-Edwards J, Willett WC. Prospective assessment of exclusive breastfeeding in relation to weight change in women. *International Journal of Obesity & Related Metabolic Disorders: Journal of the International Association for the Study of Obesity*. 2003 Jul;27(7):815–820.
- [40] Heliövaara M, Aromaa A. Parity and obesity. *J Epidemiol Community Health*. 1981 Sep;35(3):197–199.

- [41] Kaye SA, Folsom AR, Prineas RJ, Potter JD, Gapstur SM. The association of body fat distribution with lifestyle and reproductive factors in a population study of postmenopausal women. *International Journal of Obesity*. 1990 Jul;14(7):583–591.
- [42] Bjorkelund C, Lissner L, Andersson S, Lapidus L, Bengtsson C. Reproductive history in relation to relative weight and fat distribution. *International Journal of Obesity & Related Metabolic Disorders: Journal of the International Association for the Study of Obesity*. 1996 Mar;20(3):213–219.
- [43] Weng HH, Bastian LA, Taylor DH, Moser BK, Ostbye T. Number of children associated with obesity in middle-aged women and men: results from the health and retirement study. *Journal of Women's Health* (2002). 2004 Feb;13(1):85–91.
- [44] Newby PK, Dickman PW, Adami H, Wolk A. Early anthropometric measures and reproductive factors as predictors of body mass index and obesity among older women. *Int J Obes Relat Metab Disord*. 2005;29(9):1084–1092.
- [45] Bastian LA, West NA, Corcoran C, Munger RG. Number of children and the risk of obesity in older women. *Preventive Medicine*. 2005;40(1):99–104.
- [46] Ram KT, Bobby P, Hailpern SM, Lo JC, Schocken M, Skurnick J, et al. Duration of lactation is associated with lower prevalence of the metabolic syndrome in midlife–SWAN, the study of women's health across the nation. *American Journal of Obstetrics & Gynecology*. 2008 Mar;198(3):268.e1–268.e6.
- [47] Rush D, Lumey LH, Ravelli AC, Myers B. The indirect association of lactation with subsequent perimenopausal body weight. *European Journal of Clinical Nutrition*. 1996;50(1):12–16.
- [48] Bagust A, Hopkinson PK, Maslove L, Currie CJ. The projected health care burden of Type 2 diabetes in the UK from 2000 to 2060. *Diabetic Medicine*. 2002;19(s4):1–5.
- [49] González ELM, Johansson S, Wallander M, Rodríguez LAG. Trends in the prevalence and incidence of diabetes in the UK: 1996–2005. *Journal of Epidemiology and Community Health*. 2009 Apr;63(4):332–336.
- [50] Department of Health. NSF Diabetes: What is diabetes? Crown Copyright; 2010. This document, the first part of the National Service Framework for Diabetes, sets out twelve new standards and the key interventions necessary to raise the standards of diabetes care.
- [51] Manson JE, Rimm EB, Colditz GA, Stampfer MJ, Willett WC, Arky RA, et al. Parity and incidence of non-insulin-dependent diabetes mellitus. *The American Journal of Medicine*. 1992 Jul;93(1):13–18.
- [52] Hanley AJG, McKeown-Eyssen G, Harris SB, Hegele RA, Wolever TMS, Kwan J, et al. Association of parity with risk of type 2 diabetes and related metabolic disorders. *Diabetes Care*. 2002 Apr;25(4):690–695.
- [53] Nicholson WK, Asao K, Brancati F, Coresh J, Pankow JS, Powe NR. Parity and risk of type 2 diabetes: the Atherosclerosis Risk in Communities Study. *Diabetes Care*. 2006 Nov;29(11):2349–2354.

- [54] Gunderson EP, Lewis CE, Tsai A, Chiang V, Carnethon M, Quesenberry CP, et al. A 20-Year Prospective Study of Childbearing and Incidence of Diabetes in Young Women, Controlling for Glycemia Before Conception: The Coronary Artery Risk Development in Young Adults (CARDIA) Study. *Diabetes*. 2007 Dec;56(12):2990–2996.
- [55] Liu B, Jorm L, Banks E. Parity, breastfeeding, and the subsequent risk of maternal type 2 diabetes. *Diabetes Care*. 2010 Jun;33(6):1239–1241.
- [56] Stuebe AM, Rich-Edwards JW, Willett WC, Manson JE, Michels KB. Duration of Lactation and Incidence of Type 2 Diabetes. *JAMA*. 2005 Nov;294(20):2601–2610.
- [57] Villegas R, Gao Y, Yang G, Li HL, Elasy T, Zheng W, et al. Duration of breast-feeding and the incidence of type 2 diabetes mellitus in the Shanghai Women’s Health Study. *Diabetologia*. 2008 Feb;51(2):258–266.
- [58] Schwarz EB, Brown JS, Creasman JM, Stuebe A, McClure CK, Eeden SKVD, et al. Lactation and Maternal Risk of Type 2 Diabetes: A Population-based Study. *The American Journal of Medicine*. 2010;123(9):863.e1–863.e6.
- [59] Stramba-Badiale M, Fox KM, Priori SG, Collins P, Daly C, Graham I, et al. Cardiovascular diseases in women: a statement from the policy conference of the European Society of Cardiology. *Eur Heart J*. 2006 Apr;27(8):994–1005.
- [60] Lewington S, Clarke R, Qizilbash N, Peto R, Collins R. Age-specific relevance of usual blood pressure to vascular mortality: a meta-analysis of individual data for one million adults in 61 prospective studies. *Lancet*. 2002 Dec;360(9349):1903–1913.
- [61] Lewington S. The importance of cholesterol, blood pressure and smoking for coronary heart disease. *European Heart Journal*. 2003 Oct;24(19):1703–1704.
- [62] Ness RB, Harris T, Cobb J, Flegal KM, Kelsey JL, Balanger A, et al. Number of Pregnancies and the Subsequent Risk of Cardiovascular Disease. *N Engl J Med*. 1993;328(21):1528–1533.
- [63] Palmer JR, Rosenberg L, Shapiro S. Reproductive Factors and Risk of Myocardial Infarction. *Am J Epidemiol*. 1992 Aug;136(4):408–416.
- [64] Grinblatt JA, Dekker JM, Schouten EG, Palmer JR, Rosenberg L, Shapiro S, et al. Number of Pregnancies and Risk of Cardiovascular Disease. *N Engl J Med*. 1993 Dec;329(25):1893–1895.
- [65] Schwarz EB, Ray RM, Stuebe AM, Allison MA, Ness RB, Freiberg MS, et al. Duration of lactation and risk factors for maternal cardiovascular disease. *Obstetrics and Gynecology*. 2009 May;113(5):974–982.
- [66] Stuebe AM, Michels KB, Willett WC, Manson JE, Rexrode K, Rich-Edwards JW. Duration of lactation and incidence of myocardial infarction in middle to late adulthood. *American Journal of Obstetrics and Gynecology*. 2009 Feb;200(2):138.e1–8.
- [67] Vorster HH, Jerling JC, Steyn K, Badenhorst CJ, Slazus W, Venter CS, et al. Plasma fibrinogen of black South Africans: the BRISK study. *Public Health Nutrition*. 1998 Sep;1(03):169–176.

- [68] Dennison CR, Peer N, Lombard CJ, Kepe L, Levitt NS, Steyn K, et al. Cardiovascular risk and comorbid conditions among Black South Africans with hypertension in public and private primary care settings: the HiHi study. *Ethnicity & Disease*. 2007;17(3):477–483.
- [69] Levitt NS, Steyn K, Steyn NP, Lambert V, Peer N, Lombard C. A comparison of the prevalence of diabetes mellitus and associated cardiovascular risk factors in urban black South Africans;. Study Protocol Summary.

Chapter 2

The Million Women Study

2.1 Data sources and general methods

The majority of findings presented in this thesis are based on analyses of data from a large population-based cohort of middle-aged women living in England and Scotland. This study is known as the Million Women Study and was set up initially in response to concerns over the effects of hormone replacement therapy and women's risk of breast cancer and other conditions. [1] As the cohort has matured it has become possible to investigate the associations between many other risk factors and the incidence of other common diseases. This chapter will introduce the reader to the Million Women Study and will discuss the specific sources of exposure and outcomes data used in subsequent analyses. Full details of the study design and methods have been published previously. [1] Copies of the recruitment and resurvey questionnaires can be found in appendix A on page 241.

2.2 Sample

2.2.1 Recruitment

Women aged 50-64 years living in the UK are routinely invited to attend for breast cancer screening through the National Health Service Breast Cancer Screening Programme. In England and Scotland the service is run from 95 individual screening centres. [2] In 1996, NHS Breast Screening Programme centres were contacted and invited to participate in the study and 66 centres agreed to take part (64%). Figure 2.1 (page 50) shows the distribution of breast screening centres in the United Kingdom before 1998. Between 1996 and 2001 women who were invited for breast cancer screening were also invited to participate in the Million Women Study. Women were sent an initial study questionnaire¹ (designed to be brief and focused on risk factors for breast cancer) with their breast cancer screening invitation. Women who elected to participate completed the questionnaire, signed consent and returned the questionnaire to their screening centre when they attended for screening (although some women posted their questionnaires directly to the study coordinating centre). Of women who attended screening 71% returned a questionnaire. [1]

Screening centres collected study questionnaires and sent them to the study coordinating centre at the Cancer Epidemiology Unit in Oxford. Trained staff sorted study questionnaires and shredded any without appropriate signed consent (7%). Questionnaires were scanned using computerised scanning technology and data were entered into a centralised secure database.

¹Two versions of the recruitment questionnaire were used, the first version (referred to as the “aqua” questionnaire) did not contain information on breastfeeding or frequency of any physical activity, an updated version asking about these exposures (referred to as the “blue” questionnaire) was sent out after the first 9% of the study population were recruited

2.2.2 Resurvey

From 1999, postal questionnaires were sent directly to women asking them to update exposure information including medical diagnoses, treatments and procedures. Three waves of resurvey questionnaires have been sent out to date at about three, seven, and nine years after recruitment. Of the 1.3 million women who took part in the baseline survey 800,000 (65%) returned the first resurvey questionnaire, the results of which form the basis of the chapter on women's risk of reporting diabetes at follow-up. Data from the second and third waves of collection do not form part of this thesis and will not be discussed further.

2.2.3 Follow-up

Participants are routinely followed up through linkage to centrally held computerised records using their unique health care number (NHS number), date of birth and other identifying details. The linked records include NHS central registries (which contain detailed records for deaths, cancers, and emigrations), the Scottish morbidity records in Scotland (hospital admissions data), and the hospital episode statistics in England (hospital admissions data). [3, 4] The hospital admissions databases contain a record of all NHS inpatient admissions from January 1981 in Scotland and April 1997 in England. [4, 5] Diagnoses at hospital discharge (including a main diagnosis and between 5 and 13 additional diagnoses for Scotland and England respectively) or for deaths (the underlying cause of death) are coded using the World Health Organisation (WHO) classification of diseases and signs, symptoms, abnormal findings, complaints, social circumstances and external causes of injury or diseases codified in the International

Statistical Classification of Diseases and Related Health Problems 10th Revision (ICD-10). [6] Hospital and death data are complete until 31/3/2008 for women recruited in England and 31/12/2008 for women recruited in Scotland.

2.2.4 Ethics

The study was approved by the Anglia and Oxford Multi-Centre Research Ethics Committee. Access and linkage to hospital records was approved by the Information Centre for Health and Social Care in England and the Information Services Division in Scotland. All participants gave their written consent to take part in the study.

2.3 Data definitions

2.3.1 Study Population

The study population used in this thesis was restricted to women who had completed their reproductive careers. I considered a woman's reproductive career as that period of time between the menarche and the cessation of menses (i.e. menopause). I limited the study population to women who were postmenopausal at recruitment because in the Million Women Study there are reasonably few women who are pre- or perimenopausal at recruitment (4.4%), being postmenopausal is an objective measure of completed childbearing (i.e. not self-reported completed childbearing), women are not naturally able to change their childbearing exposures postmenopausally, it is difficult to accurately classify women who are pre- and peri-menopausal (especially in the absence of biological data.)

Menopause

I used the standard Cancer Epidemiology Unit Million Women Study data definition to classify women as postmenopausal if they reported their periods stopped naturally or after bilateral oophorectomy, or if they were aged 53 years or older (most women aged 53 years who had a natural menopause were postmenopausal by that age).

2.3.2 Exposures

Parity

At recruitment women were asked “Have you ever had any children?” and if yes, “How many children have you had? (please include stillbirths; it is not necessary to include miscarriages)” this information was used to define women’s parity. In keeping with Cancer Epidemiology Unit standard practice I defined parity using the number of full term pregnancies a woman self-reported (this includes still births). For women with multiple pregnancies a single parous event is recorded. (Diana Bull, personal communication).

Figure 2.2 on page 50 shows the distribution of parity within the study population. About 11% of women were nulliparous, 13%, 41%, 22%, and 8% of women in this study reported giving birth to one, two, three or four children respectively. Less than 5% of all the women in the study had given birth to more than four children. As a result I summarised parity into 0, 1, 2, 3, or 4 or more children. For women with a parity of four or more the mean parity was 4.4 [SD 0.8].

Breastfeeding

After the first 9% of the study population were recruited, questions on women's history of breastfeeding were added to the study questionnaire ². Women were asked, "How many months did you breastfeed each child, if at all?" Figure 2.3 on page 51 shows the distribution of breastfeeding duration in months among parous women who ever breastfed. No questions were asked about the exclusivity of breastfeeding. It can be seen from figure 2.3 that there is evidence of digit preference at 6, 12, 18, and 24 months.

I used information from the recruitment survey to define breastfeeding (ever or never), total duration of breastfeeding (summation over all children of reported duration of breastfeeding in months), and duration of breastfeeding per child (parity/total duration of breastfeeding) ³ Duration of breastfeeding was generally included in various statistical models as a categorical variable.

I used information from the recruitment survey to define breastfeeding (ever or never), total duration of breastfeeding (summation over all children of reported duration of breastfeeding in months), and duration of breastfeeding per child (parity/total duration of breastfeeding). Duration of breastfeeding was generally included in various statistical models as a categorical variable.⁴

²Two versions of the recruitment questionnaire were used, the first version (referred to as the "aqua" questionnaire) did not contain information on breastfeeding or frequency of any physical activity, an updated version asking about these exposures (referred to as the "blue" questionnaire) was sent out after the first 9% of the study population were recruited

³I explored creating a "per child duration of breastfeeding" variable using women's parity as the denominator and alternatively the number of children breastfed as the denominator. For example, for a woman with a parity of 4 who breastfed 3 of her children with a total duration of breastfeeding of 12 months the average "per child duration of breastfeeding" would be either 3 months per child using parity as the denominator, or 4 months per child using the number of children breastfed. Parity and the number of children breastfed were highly correlated (Spearman's correlation coefficient 0.65, $p < 0.00001$). I comparing the methods of calculating the "per child duration of breastfeeding" using the means and distribution of the data. The average duration of breastfeeding per child differed by less than half-a-month between the two methods (mean difference 0.47 [SE 0.001]). I therefore decided to define per child duration of breastfeeding using parity as the denominator.

⁴Women's reproductive data are consolidated prior to being made available as part of the Million Women Study database. For women with singleton pregnancies parity and duration of breastfeeding are

In chapter 3 (page 63) breastfeeding duration data from the Million Women Study is compared with breastfeeding duration data from the 1958 and 1970 birth cohorts. However, there were no standard categories for the duration of breastfeeding that could be applied across all studies due to the use of different measures of the duration for which women breastfed the index child. In order to facilitate direct comparison between data from the Million Women Study and the two birth cohorts I created standard measures of the duration of breastfeeding in the Million Women Study using the definition appropriate for each reference study. Detailed definitions of duration of breastfeeding used in this analysis can be found in chapter 3 (page 63)

In general, for the analyses of the association between childbearing and BMI, risk of diabetes mellitus, or risk of IHD women's total duration of breastfeeding was categorised as never breastfed, breastfed for less than 6 months, breastfed for 6 to 9 months, or breastfed for 10 months or more (which approximated quartiles of total duration for women who breastfed). The means of these categories for women who breastfed were 2.3 months [SD 1.3] for women who breastfed less than 6 months, 7.2 months [SD 1.2] for women who breastfed between 6 and 9 months, and 18.3 months [SD 11] for women reporting a total duration of breastfeeding of 10 or more months. It should be noted in figure 4.6 (page 114) presented in chapter 4, I used a summary of the duration of breastfeeding per child (< 6 months, ≥ 6 months) to illustrate some of my findings.

defined as above. For women with multiple pregnancies each pregnancy is considered a single parous event, however each child breastfed contributes to the total duration of breastfeeding.

2.3.3 Outcomes

Body Mass Index (BMI)

At recruitment, women were asked to report their current weight and height. I used these measures to derive mean body mass index (BMI) as $\text{weight (kg)} / \text{height (m)}^2$. Women's height and weight data were restricted to biologically plausible ranges, 120 to 200 cm for height and 35 to 195 kgs for weight. For analyses presented in chapter 4 (page 87) BMI is the primary outcome and was modeled as a continuous variable.

Diabetes Mellitus

At recruitment, women were asked "Have you ever had [Diabetes]?" and "Are you now being treated for [Diabetes]?" At the first repeat survey around 3 years later women were asked "Have you had any of the following conditions [Diabetes] diagnosed for the first time in the last 5 years?", and "Are you now being treated for: [Diabetes] If yes, about how old were you when treatment began?" The outcome of interest analysed in this thesis was women's self-report of a diagnosis of, or treatment for, diabetes mellitus at resurvey and which occurred after the date of recruitment in women with no prior history of diabetes mellitus.

Ischaemic heart disease

For analyses in this thesis I defined ischaemic heart disease (IHD) as either the primary hospital discharge diagnosis or death (due to) based on the International Statistical Classification of Diseases and Related Health Problems 10th Revision (ICD-10) codes I20 to I25 (inclusive) found in Chapter IX: Diseases of the circulatory system. [6]

2.3.4 Covariates

In conceptualising the associations between childbearing and BMI and risk of diabetes and heart disease I began by considering the associations between childbearing and BMI (BMI is a known risk factor for both diabetes and heart disease so it was important to explore what association might exist between childbearing and BMI in order to appropriately consider how to manage BMI in the analyses of the associations between childbearing and these outcomes.)

I considered that at a given age one's BMI is a combination of remote and recent influences. For example, in women who have completed their childbearing years BMI is likely a combination remote factors like post-partum BMI (for each parity event) and more recent BMI-influencing factors (like current levels physical activity.) Age, region of residence, socioeconomic group, and smoking were important co-factors (potential confounders) I included in all of the analyses. Where appropriate the decision to include other co-factors including physical activity and alcohol was based on findings from exploratory analyses of the co-factor (potential confounder) in question⁵

In conceptualising the associations between childbearing and women's risk of diabetes and heart disease I considered that BMI would likely be an important mediating factor (and part of the causal pathway.) I explored the associations between childbearing and risk of diabetes and heart disease with and without BMI to assess the directness of associations between childbearing and women's risk of diabetes and heart disease.

Below are definitions of the various co- and mediating factors I considered in my analyses.

⁵One of the factors influencing post-partum BMI is pre-pregnancy BMI although I had access to a crude measure of pre-pregnancy body size (clothes size at age 20 years, average age at first birth in the study population was 24 years) which I could use in sensitivity analyses ideally I would have liked to have access to a good measure of pre-pregnancy BMI which I could consider in various models.

2.3.4.1 Defining Co-factors

Age - the Million Women Study recruited women aged 50 to 64 years. At recruitment women were asked to record either their date of birth or their current age or both. In the analyses of the association between childbearing and BMI which used linear regression models, age at recruitment was included as a continuous variable. For the analyses of the association between childbearing and risk of diabetes mellitus using logistic regression models, attained age (age at repeat first survey) was included as a continuous variable. In the analyses of the association between childbearing and risk of ischaemic heart disease based on Cox regression modelling, attained age was used as the underlying time variable. In addition it was included in the model as an explanatory variable (summarised into 5 year categories - <50, 50 to 54.9, 55 to 59.9, 60 to 64.9, 65 to 69.9, 70 years or older).

Region of residence - participants in the Million Women Study are classified into 10 geographic regions within England and Scotland. Region of residence is a derived variable based on the geographic location of the breast screening centre and cancer registry area where a woman was registered. I included region as a categorical variable in all models.

Socioeconomic group - using the postcode of their usual residential address as reported at recruitment, women were assigned to an enumeration district (England) or an organisational area (Scotland). This information was then converted into a deprivation score based on the Townsend index of deprivation. [7] Women were categorised on the basis of their deprivation score into fifths. In addition, for some analyses I derived a summary variable for socioeconomic group based on tertiles of deprivation.

Smoking - women were asked about their smoking habits at recruitment and at resurvey. At recruitment women were asked “About how many cigarettes do you smoke on average each day, now? none, less than 5 per day, 5 - 9, 10 - 14, 15 - 19, 20 - 24, 25 or more”, at the first repeat survey about 3 years later, women were asked for more detailed information on their past and current tobacco use including the questions “Have you ever been a smoker?”, and “About how many cigarettes do you/did you smoke on average each day? (if you are an ex-smoker, how many did you smoke on average when you smoked?)”

Smoking is known to affect the risk for each of the outcomes analysed in this thesis, [8–11] I therefore included it as a co-factor in all of the analyses. Based on information at recruitment women were categorised as never, past, or current smokers. For current smokers, the number of cigarettes smoked per day were broadly grouped into <15 cigarettes per day, or ≥ 15 cigarettes per day for the univariate analyses, while for the multivariate analyses finer categories of the number of cigarettes were used; 1 to 4, 5 to 9, 10 to 14, 15 to 19, 20 to 24, or ≥ 25 . Where appropriate I addressed the potential for misclassification of exposure levels using the regression dilution approach. [12, 13]

Physical activity - at recruitment women were asked to report on the frequency and intensity of physical activity. For the analysis of the effects of childbearing on women’s mean BMI (chapter 4, page 87), I derived a summary variable for physical activity using self-report of any physical activity (“How often do you do *any* exercise?”). For the analyses of the effects of childbearing on women’s risk of diabetes (chapter 5, page 122) and ischaemic heart disease (chapter 6, page 168) I also derived a summary variable using women’s response to the strenuous physical activity question, which was defined

as exercise that is “enough to cause sweating or a fast heart beat.” For the appropriate univariate analyses, physical activity and strenuous physical activity were broadly grouped into “ever” versus “never”. For the multivariate analyses finer categories of exposure were used. For example, I summarised “any physical activity” into three categories; ≤ 1 time per week, 2 to 3 times per week, ≥ 4 times per week. “Strenuous physical activity” was summarised into four categories; none, ≤ 1 time per week, 2 to 3 times per week, 4 or more times per week.

Alcohol - women were asked at recruitment to report on average, each week, how much wine (red, white, or both), beer, and spirits they drank. Women were asked again about their alcohol intake (frequency and type) at the first repeat survey (about 3 years later). Reported intake was converted into units of alcohol (one standard drink being equal to one unit). Weekly alcohol intake was summarised into categories as follows; none, 1 to 6, 7 to 14, or ≥ 15 units per week. I also converted reported intake into grams of alcohol consumed (one unit was considered equivalent to 10g of alcohol). Again, I addressed the potential for misclassification of exposure levels using the regression dilution approach. [12, 13]

Please see the data analyses chapters for discussions on other co-factors which were considered, and for additional descriptions on how co-factors were managed. In addition, where previous published studies had considered other co-factors I performed additional analyses to explore for similar effects in Million Women Study data.

2.3.4.2 Defining mediating Factors

Body mass index (BMI) As well as being an outcome BMI was also considered to be an important potential mediating factor between childbearing and diabetes and heart

disease. I explored associations between childbearing and diabetes and heart disease accounting for potential mediating effects of BMI. As mentioned earlier in this chapter (section 2.3.3, page 25) BMI was defined as $\text{weight (kg)} / \text{height (m)}^2$ based on self-reported height and weight data obtained at recruitment. For analyses presented in chapters 5 and 6; pages 122, 168) I modelled BMI as a continuous or categorical variable as appropriate. In general, where BMI was used as a categorical variable it was divided into the following categories; <20 , 20 to 22.4, 22.5 to 24.9, 25 to 27.4, 27.5 to 29.9, 30 to 32.4, 32.5 to 34.9, or $\geq 35 \text{ kg/m}^2$, and I used 22.5 to 24.9 kg/m^2 as the reference category. In chapter 5 (page 122) I presented stratum specific results of the association between childbearing and diabetes mellitus using the following BMI categories <25 , 25 to 30, 30 to 35, or $\geq 35 \text{ kg/m}^2$. I also considered the potential effects of misclassification of BMI using the regression dilution approach (this is discussed in section 2.5 on page 47). [12, 13]

2.4 Data quality

2.4.1 Sample

2.4.1.1 Data entry & storage

Trained staff sort, check, and electronically scan all the study questionnaires. [1] Software that makes use of computerized intelligent character recognition is used to capture scanned data. [Eyes and Hands ®, Readsoft Inc, Slough, UK] Range and logical checks are performed at the time of data entry and trained staff verify inconsistencies and any information that is not recognised by the data capture software. Staff confirm consent for follow-up has been granted and verify computer-interpreted data. From

each day's work a random sample of questionnaires is drawn and the scanned and verified data are checked against the original questionnaires, the error rate is consistently <1% (personal communication Mrs Barbara Crossley).

Personal identifying information is removed before data are entered into a secure database. Access to the database is restricted to named research staff. Trained database administrators regularly up date the database as new data becomes available. Database administrators prepare specific data views for approved research projects.

2.4.1.2 Non-response & generalisability

Eligible women were recruited into the study via the NHS Breast Cancer Screening Programme. Pilot studies carried out before the main study began showed that recruitment into the study did not affect screening up-take. [14] An analysis of women who did and did not attend breast cancer screening found that women who attended for screening were similar in age and general medical profile to women who did not attend but were less deprived and more likely to have been prescribed hormone replacement therapy. [15] The characteristics of the study population have been published previously. [1] Briefly, women who returned a questionnaire and provided signed consent for follow-up were slightly more likely to have used oral contraceptives and to be current users of hormone replacement therapy than those who did not.

All of the women who were recruited into the Million Women Study were sent a copy of the first repeat survey questionnaire. Of these 866,551 (63%) returned a completed questionnaire. Questionnaire response rates were higher for women from regions in the south of England, of higher socioeconomic status, with lower body mass index,

who regularly consumed alcohol, and who were never or past smokers. These results are shown and discussed in more detail in chapter 5 (page 122).

2.4.2 Exposures

Exposure data analysed in this thesis are based on women's recalled childbearing histories which may be affected by bias and misclassification. As part of the preliminary work for this project I explored the extent to which recalled parity and breastfeeding are *valid* and *reliable*. By *valid* I mean "the independent corroboration that the instrument is measuring what it intends to measure." (in the literature this is a subset of *Criterion Validity*). [16] By *reliable* I mean the consistency (repeatability) of participants responses to the same (or similar) questions asked at two or more time-points. [17]

2.4.2.1 Parity

A comprehensive search of the current literature revealed several studies of the *reliability* of women's recall of their parity but none assessed the *validity* of the measure⁶. [18–25] *Reliability* was assessed by asking women to self-report number of previous births (most often live births) at recruitment and then again at resurvey and comparing the agreement between the data sources (test-retest repeatability). The studies varied in size from 46 to 535 mothers, and the test-retest interval ranged from 8 weeks to 25 years. Data were collected through self-administered questionnaires, interview-administered questionnaires (in-person or over the telephone), and medical records. Women were asked to report on the number of pregnancies, number of live births or both.

⁶validity of parity has been assessed for single pregnancies, but I could not find any published literature the validity of women's recalled parity once her reproductive years have been completed

Table 2.1 (page 52) summarises the relevant published findings. All of the studies reported *reliability* using correlation methodology. Four of the studies reported the correlations between questionnaires for number of live births which were between 0.88 and 1.00. [18, 20–22] Two studies reported the *reliability* of levels of parity (defined as “live births”. In a small study (n=53) of the mothers of offspring who had participated in the National Collaborative Perinatal Project (psychotic offspring and matched controls) Buka et al tested the *reliability* of a “Pregnancy History Instrument”; investigators found that for cases the correlation for parity (at the time of the initial study questionnaire) 0 or parity ≥ 4 was 1, for controls the correlation for parity 0 was 1 while for parity ≥ 4 it was 0.89 (this is likely an effect of the small numbers, only one woman had conflicting data). [19] In a much larger study (n=535) of the *reliability* of reproductive histories reported by a population of postmenopausal women Lucas et al found high overall percentage agreement between interview administered questionnaires (94.8%) and a correspondingly high correlation coefficient (kappa = 0.93). Investigators also reported the percentage agreement for individual parity levels. It can be seen from the table 2.1 (page 52) that as parity increased the percentage agreement decreased. Two studies reported the correlations between questions asked about the mean number of pregnancies; Slanger et al tested the repeatability of a questionnaire of women’s reproductive histories through two in-person interviews about 40 weeks apart in a random sample of women drawn from a large population-based case-control study of the associations between hormone replacement therapy and breast cancer risk. Investigators reported the percentage agreement for pregnancies was 91% and the kappa was 0.92. [23] In contrast, Ko et al found a high percentage agreement for pregnancy history (99.7%) but a lower correlation (kappa = 0.67) in a study of the repeatability of a reproductive history questionnaire administered by interview with a test-retest interval of 8

years. [24]

In addition to reviewing the published literature base I also explored the reliability of childbearing data in a sub-group of women who answered the recruitment questionnaire twice. I found that Million Women Study participants reported their childbearing histories reliably (see Appendix B, page 250.)

2.4.2.2 Breastfeeding

The published evidence for the *validity* and *reliability* of long-term maternal recall of breastfeeding was very limited. I found only four studies exploring *validity* and five exploring *reliability* of long-term maternal recall of breastfeeding. Findings from the literature are summarised in tables 2.4 and 2.5 (pages 55, 56).

The studies exploring the *validity* of long-term maternal recall of breastfeeding were small (sample sizes between 74 and 144 mothers) and largely depended on the serendipitous availability of data. For example in the Jerusalem-based study by Kark et al, mothers of military recruits (aged 20 to 22 years) were asked to recall how they fed the index child. This data were then compared with available clinic records. 8,646 participants were enrolled in the study (average age 17 years) but clinic records “had been destroyed in all but two of the numerous clinics in the city ”and records were only available for 125 subjects. The validation study was based on the 74 participants with complete data. [26] Whilst in one of the largest of the validation studies (144 women) researchers used prospective breastfeeding data collected fortuitously during the course of another study – the Menstruation and Reproductive History Study and compared it to recalled data collected between 35 and 50 years later. [27]

Despite these limitations study investigators generally found that as compared to prospectively collected data, long-term maternal recall was a valid measure of women's breastfeeding histories. For example in Kark et al which compared maternal recall of breastfeeding after 17 years with matched post-natal clinic records, investigators found high overall correlation between the two data sources (Spearman's correlation 0.86). In addition when breastfeeding was split into short (≤ 5 months) or longer (> 5 months) durations the percentage agreement was 91% and 85% respectively. [26] Similarly, as part of a study of Australian adolescents, Tienboon et al asked mothers of 144 adolescents to recall how they fed the index child and compared this information with available data from infant welfare cards. There was high agreement between the data sources for ever breastfeeding (percentage agreement 85%), even though investigators noted the agreement was better for first and second born children than for later born children. For women who reported having ever breastfed the correlation between recorded and recalled breastfeeding duration was 0.77. Investigators reported that women tended to overestimate the duration they had breastfed (recalled duration of breastfeeding 6 months (SD 3.5), recorded duration of breastfeeding 4.7 months (SD 3.6)). [28] In a study of 95 low middle-class Canadian women Vobecky et al found a high correlation between duration of breastfeeding data collected prospectively and duration of breastfeeding data recalled 8 years later (correlation coefficient 0.93). [29] In a study of 140 US postmenopausal women Promislow et al reported that maternal recall of ever having breastfed had high sensitivity (94%). Although only 26% of women had perfect agreement between recorded and recalled durations of breastfeeding expressed in months, 71% of women correctly recalled the duration of breastfeeding within two months of duration reported in the recorded data. The overall weighted kappa statistic for the

association between the two sources of data were 0.55 (95%CI 0.42 to 0.67). Digit preference at 9 and 12 months was noted. [27]

The published literature examining the *reliability* of long-term maternal recall of breastfeeding contained few generally small studies. [18, 22, 24, 26–30] *Reliability* was assessed by asking parous women to report if they had breastfed and for how long on two occasions and comparing the agreement between the data sources. All of the studies reported test-retest repeatability using correlation methodology. The studies included between 107 and 564 mothers and the test-retest interval ranged from 8 months to 8 years. Again, although the evidence base was limited, investigators generally reported that women *reliably* recall their histories of breastfeeding even with a prolonged interval between questionnaires. All of the studies reported high correlation coefficients (0.86 to 0.89) except for one recent study among middle-aged Korean women. Ko et al explored the repeatability of a reproductive history questionnaire administered by interview with a test-retest interval of 8 years in 386 Korean women; the study investigators found that although the percentage agreement between questionnaires of having ever breastfed was high (96%) the kappa was very low (kappa 0.20). [24] This finding may be due in part to the very high prevalence of breastfeeding in this population (95%). Of the 9 women who reported never breastfeeding on at least one occasion only two reported never breastfeeding on both occasions. [31]

2.4.3 Outcomes

Central to the success of a cohort study is the study team's ability to trace participants and to capture, as completely as possible, deaths, emigrations and outcomes of interest. [17, 32, 33] Investigators from the Million Women Study have access to and use

recruitment data and repeat survey data (which contains update exposure information and records of outcomes of interest). Study investigators have also linked participant records to central registries (Office of National Statistics) to obtain detailed information on deaths, emigrations and cancers that occur during follow-up, and to medical records, primarily hospital admissions data from Scotland (Scottish Morbidity Record) and from England (Hospital Episode Statistics, HES), which has allowed study investigators to track non-fatal, non-cancerous outcomes of interest such as ischaemic heart disease. Ideally study participants would also be linked to outpatient and general practice records so that illnesses that do not require hospital admission, for example diagnosis and treatment of uncomplicated diabetes mellitus, could be captured.

The linkage of study participants to these datasets and the assessment of data quality have been the primary focus of DPhil theses by Dr Xian Sweetland (University of Oxford) and Dr Bette Liu (University of Oxford). As such these aspects will not be discussed in detail. Work undertaken to assess aspects of data quality that have not been covered in these theses are discussed in detail in sections 2.4 on page 30.

2.4.3.1 Body mass index (BMI)

It is known that BMI calculated from self-report is often under estimated due to people slightly over reporting their heights and under reporting their weights. [34] In 2008 study investigators invited selected participants from the Million Women Study to take part in a height, weight, waist and hip measurement validation study. Women and their associated GP practices who agreed to participate were sent a data collection kit which included instructions for the practice nurse or GP, a tape measure, and a form to record the measurements. Women were asked to send the kit and the measurement forms back to the study coordinating centre. (Dr Lucy Wright - personal communication)

For the random sample of 3,745 women who participated in the validation study there were three sources of BMI data available; BMI calculated using self-reported heights and weights from the recruitment questionnaire, BMI calculated from self-reported height and weight data collected about 8 years after recruitment (for the validation study), and BMI calculated using measured heights and weights (obtained within 6 months of the validation study survey).

Using data from the validation study, and in consultation with Dr Lucy Wright and Dr Ben Cairns, I explored the degree of and possible effects of misclassification of BMI. Firstly, I compared BMIs calculated from self-reported height and weight data with BMIs calculated from measured height and weight data using correlation and Bland-Altman techniques. [31, 35, 36] As will be discussed more fully in chapter 4 (page 87) the correlation between the two data sources was very high (Spearman's correlation coefficient 0.95). The Bland-Altman analysis suggested a slight degree of underestimation of BMIs from self-reported data as compared to BMIs from calculated data; however the mean difference between the two measures was not significantly different to zero ($-0.65 \text{ kg}/\text{m}^2$, 95% CI -3.10 to $1.79 \text{ kg}/\text{m}^2$).

To explore the effects that misclassification of BMI might have on the findings I compared the model estimates obtained using BMI data from self-report with model estimates obtained using measured BMI data. For example, in a sensitivity analysis performed as part of the analyses of the association between women's childbearing histories and their postmenopausal BMI I restricted the study population to women with measured BMI data.

2.4.3.2 Diabetes Mellitus

There is an established evidence base showing the *validity* of self-reported diabetes. [20, 37–42] It has also been shown in this study population that women's report of treatment for diabetes strongly agrees with general practitioner prescription records for treatment of diabetes (percentage agreement 98%, kappa 0.78). [43] Investigators from the EPIC-Potsdam study and from the British Regional Heart Study compared how reliably participants self-reported a diagnosis of diabetes on two separate occasions where the recall interval was two years and five years respectively. [44, 45] Both studies compared self-reported medical histories obtained from an in-person interview with those from a self-administered questionnaire. Participants reliably recalled a diagnosis of diabetes; for participants in the EPIC-Potsdam study the percentage agreement was 74% with a kappa of 0.84. For participants in the British Regional Heart Study the percentage agreement was 84%. Table 2.6 (page 57) shows for women participating in the Million Women Study the agreement between women's self-reported diabetes at recruitment and at the first repeat survey. 83% of women who reported diabetes at recruitment also reported diabetes at the first repeat survey. Findings from the Million Women Study on the *validity* and *reliability* of self-reported diabetes are comparable to the published findings from other large European population-based studies. For analyses from this study population presented in this thesis women's self-report of diabetes is an acceptably *valid* and *reliable* measure.

2.4.3.3 Ischaemic heart disease

As was described in section chapter 2.2.3 (page 20), study participants have been linked to the NHS central registries, using a combination of unique identifying details (name,

NHS number, date of birth, etc.), which contain detailed records for deaths, emigrations, and cancers.

English and Scottish mortality data are provided by the Registrar General and the General Register Office for Scotland respectively. [46, 47] For each death the conditions listed on the death certificate are used, according to ICD rules, to specify the “underlying cause of death”⁷. The quality of death certification data are routinely checked through a hierarchical process of automated checks in the registration and cause of death coding processes. [46] In addition the Automated Cause Coding System is regularly checked by comparing results with a randomly selected sample (10%) of manually coded records. The error rate is “extremely low” (personal communication D.Goodwin).

Morbidity data on ischaemic heart disease and stroke used in this thesis are taken from administrative databases of morbidity in Scotland (Scottish Morbidity Records) and England (Hospital Episode Statistics). [4, 5] In Scotland acute inpatient or day case morbidity data has been systematically collected through the Scottish Morbidity Record (SMR) system and held in a linked database by the Information and Statistics Division of the Scottish Executive (ISD) since 1981. ISD Data Quality Assessment staff (DQA) are responsible for auditing the SMR records. DQA staff visit each hospital and compare data from sample of SMR records with medical case notes. The ISD Data Quality Assurance Report on Acute Inpatient/ Day Case Data 2000 – 2002 reported accuracy for main clinical condition [primary diagnosis] of 88%. [5]

In England morbidity data for inpatient care from the National Health Service has been routinely collected since 1989 through the Hospital Episode Statistics (HES) which are

⁷World Health Organisation defines the underlying cause of death as the disease or injury that initiated the train of events directly leading to death; or the circumstances of the accident or violence that produced the fatal injury.

generated from Secondary Use Service. [4] HES data are internally audited, in addition the quality is assessed using the Data Quality Indicator (DQI) which evaluates the data coverage and logical consistency. Reports on the completeness and quality of HES data are produced monthly and annually. The DQI summary report for 2002/2003 noted a composite DQI component score of 93.4% which suggests that for the conditions audited the data are of good quality. [4] However, ischaemic heart disease is not one of the routinely audited conditions and there is little published evidence on the quality of HES data for this outcome. The Oxford Vascular Study which compared all acute vascular events in the Oxfordshire region with HES data found that 79% of study participants known to have IHD had a hospital discharge diagnosis consistent with this (ICD-10 I20-I25). [48]

2.4.4 Co-factors

2.4.4.1 Smoking

From the published evidence self-reported smoking status and intensity are known to be reasonably well reported. [10, 49–51] Additional work on how well questions from the Million Women Study questionnaires capture women's smoking habits is currently under way and will not be discussed further in this thesis (personal communication Dr B. Cairns).

2.4.4.2 Physical activity

In general, reasonable evidence exists for the reliability and validity of self-reported data on physical activity. [52] In the Million Women Study women were asked at recruitment “How often do you do any exercise?”, and “How often do you do any strenuous exercise? (that is, enough to cause sweating or a fast heart beat?)” The repeatability of this information was assessed using data from the 19,059 (1.5%) of the 1.3 million participants who completed the recruitment questionnaire twice⁸. Study investigators found moderate degrees of agreement for these simple survey based measures of physical activity (the percentage agreement and kappa were 57% and 0.42, and 63% and 0.42 for any physical activity and strenuous physical activity respectively). Investigators also reported a small but progressive decline in agreement was found with increasing time between questionnaires. [53] Given these findings it is not possible to rule out some misclassification of levels of physical activity.

2.4.4.3 Alcohol

Reported alcohol intake has been shown to have good reproducibility in several population-based studies. [54, 55] For example, as part of the European Prospective Investigation into Cancer and Nutrition (EPIC) a small validation study which included 58 Dutch women compared information on nutrient intake obtained from a standard Food Frequency Questionnaire (FFQ) at the start of the study with monthly 24-hour food-intake recall questionnaires (one per month for 12 months). [55] The Spearman rank correlation coefficient for the two methods was high (0.87). Women also reliably reported

⁸Women who had inadvertently been recruited twice. For example if as a result of moving house a woman moved from one screening centre to another and received, completed, and returned two recruitment questionnaires

their alcohol intake (correlation coefficients comparing the food frequency questionnaire with a 24-hour dietary recall at six months and at 12 months were 0.91 and 0.92 respectively). Previously published work has shown that among Million Women Study participants reported alcohol intake is reasonably reliable (energy adjusted weighted kappa = 0.59). [56]

Please see the data analyses chapters for discussions on other co-factors which were considered, and for additional descriptions on how co-factors were managed.

2.4.5 Exclusions and missing data

In this thesis exclusions either on an *a-priori* basis (for example pre-menopausal women) or due to incomplete or missing data, were applied in a hierarchical manner. The exact numbers of women excluded can be found in the methods section of the relevant data analysis chapter. Below are brief descriptions of the *a-priori* exclusion criteria applied in this thesis and how missing or incomplete data were coded and handled in the analyses.

2.4.5.1 Exclusions

Women younger than 49 years (381, 0.03%) or older than 65 years (28,478, 2%) at recruitment were excluded from all analyses as were women missing information on age (50, <0.01%). I used the menopause as the natural boundary signalling the end of women's reproductive years. For the purposes of this thesis I restricted our eligible study population to postmenopausal women. Women were classified as postmenopausal if they reported their periods stopped naturally or after bilateral oophorectomy, or if they were aged 53 years or older - most of the cohort aged 53 years whose menopausal status was

not masked (for example by HRT) were postmenopausal by that age (personal communication Diana Bull).

For the analyses looking at the effects of childbearing on BMI and risk of diabetes I excluded women who had been registered with a cancer (other than non-melanoma skin cancer) before recruitment. I decided to exclude these women because previous cancer and cancer treatments are known to affect BMI and could reasonably be expected to affect the risk of diabetes mellitus.

2.4.5.2 Missing data

Exposures

Women with missing information on the exposures of interest were excluded from all the analyses.

Parity - women who left the question, "Have you ever had any children?" blank or who answered "yes" but did not give any further details on the number of children they had given birth to were coded as missing.

Breastfeeding - a question on breastfeeding ("How many months did you breastfeed each child, if at all?") was added to the questionnaire after about the first 9% of the study population had been recruited. Women who had not been asked about breastfeeding were excluded. In addition parous women who had been asked about breastfeeding but for whom the data were either missing or incomplete were also excluded. I explored potential patterns of missingness for breastfeeding by selected cofactors including parity, socioeconomic group, and smoking. I also explored whether certain

co-factors predict missingness of breastfeeding information. The only covariate that was both a meaningful and significant predictor of missingness was nulliparity.

Outcomes

Women missing outcome data were managed according to the type of analyses I performed.

BMI - I calculated mean BMI from women's self-reported height and weight at recruitment. Women missing data for either of these variables were coded as missing for BMI. In addition I coded as missing women with either a height or weight outside a biologically plausible range (< 120 cm or > 210 cm for height , < 35 kg or > 195 kg for weight). For the cross-sectional analysis of the effects of childbearing on women's mean BMI, I excluded women with missing BMI. In analyses where BMI was a co-factor I included women with a missing BMI in a separate category.

Diabetes - I had complete outcomes data for women who were eligible for inclusion in the analyses on the effects of childbearing on women's risk of diabetes.

IHD - for the analyses of the effects of childbearing on women's risk of ischaemic heart disease I right censored the data as appropriate. Women contributed person-years from recruitment until the date of their first IHD event, death, emigration or end of follow-up, whichever was earliest.

Co-factors

Study participants with missing values recorded for co-factors were managed in two ways. For the analysis of the main effects of childbearing on women's mean BMI (chapter 4, page 87) the main analysis was restricted to women with no missing values for any of the co-factors included in the model. For analyses of the effects of childbearing on women's risk of diabetes (chapter 5, page 122) and ischaemic heart disease (chapter 6, page 168), women with missing values for any of the adjustment variables were assigned to a separate category. Sensitivity analyses were subsequently carried out restricting the analyses to women without missing values. Below are brief descriptions of which co-factor categories were coded as missing. Additional details can be found in the relevant data analysis chapters.

Socioeconomic group - socioeconomic group was derived using the Townsend deprivation index. [7] Deprivation was coded as missing where the postcode was missing or incomplete, where the enumeration district or organisational area was found but there was no corresponding deprivation index given, where the postcode was valid but there was no enumeration district (at that time - 1991), or where the postcode appeared valid but was not found on the census database.

Smoking - at recruitment women were asked questions on smoking (see 2.3.4.1, page 28) including "Are you an ex-smoker?" Women who left both questions blank can correctly be said to have missing data on their smoking exposure. In addition women who could not be correctly classified as either ex- or never smokers, because they reported not currently smoking any cigarettes but left the the ex-smoker question blank and could be either ex- or never smokers, were also classified as missing.

Physical activity - women were asked about frequency and intensity of strenuous physical activity in both versions of the recruitment questionnaire. Women who left this question blank had physical activity coded as missing. In addition a question on the frequency and intensity of any physical activity was added to the questionnaire after the first 9% of the study population had been recruited. Women had data on any physical activity coded as missing if they had not been asked the question or if they left the question field blank.

Alcohol - using information from the recruitment questionnaire women who left all of the questions on the type or frequency of alcohol intake were coded as missing. In keeping with established unit procedures I calculated alcohol intake for women who answered any question on type of alcohol consumed assuming that for types of alcohol left blank the average intake was zero.

2.5 Data analysis

In this thesis data analyses were conducted using the STATA statistical software package (Stata corporation, TX, USA). All of the data were analysed using standard statistical methods and techniques. Where appropriate, data were described using the mean and standard deviation for continuous variables and frequencies (percentages) for categorical variables. For the analyses where the exposure and outcome were measured at a single time point (either at the same time in the case of the analyses quantifying the association between childbearing and women's body mass index, or at two different time points as in the analyses quantifying the association between childbearing

and women's risk of diabetes after the menopause) linear and logistic regression models were used. For the analysis quantifying the association between childbearing and women's risk of IHD after the menopause the outcome event (IHD) occurred over a variable period of follow-up time I made use of Cox-regression models with attained age as the underlying time variable. The risks for diabetes or IHD associated with childbearing are presented as floated odds ratios and floated hazard ratios respectively. Floating the absolute risk adds variance to the reference category thereby reducing the variance of other categories without altering the risk estimates. [57] Where appropriate I used the category means to performed trend tests. In order to account for possible misclassification of certain exposures, for example smoking or alcohol intake, I summarised each exposure of interest into categories based on information from the recruitment questionnaire, however the average for each category was calculated using more recently obtained data (for this thesis data as reported on the repeat survey questionnaire administered about 3 years after recruitment from 708,265 women). This technique is described in the literature as the regression dilution approach. [12, 13] Details of more specific analyses are explained in the methods section of each of the following chapters.

2.6 Discussion

The Million Women Study is a very large prospective cohort study of middle-aged women living in England and Scotland. The study was set up initially in response to concerns over the effects of hormone replacement therapy on women's health and risk of disease and as the cohort has matured it has become possible to investigate the associations between many risk factors and the incidence of common diseases. As the

Million Women Study is a major source of the data used in this thesis, it is important to have an in-depth understanding of the available data. As part of this process I reviewed the study methods used at recruitment, repeat survey, and follow-up of participants (including appropriate ethics approval). I explored the general quality of the data and specifically the quality of parity, and breastfeeding data including a comprehensive review of the current literature. I also explored and commented on the quality of the data available for our primary outcomes and for important co-factors. I found and reported the evidence that the data to be used in analyses including the long-term recall of childbearing information, self-reported height and weight data used to calculate BMI, self-reported diabetes, IHD diagnosis from hospital admissions databases, and on important co-factors were of good quality. In addition, sufficient supplementary data are available to allow for validation studies e.g. example measured height and weight data. Data from the Million Women Study used in this thesis are of sufficient quality to allow for meaningful analyses and cautious interpretation of results for the exposures and outcomes specified.

2.7 Chapter summary

In this chapter I have introduced and discussed the Million Women Study describing how the study came about and how the study sample was obtained. I introduced and defined the exposures, outcomes, and important co-factors to be used in the subsequent analyses. I explored issues around the quality of the data and showed that the data are of a high quality and suitable for the planned analyses.



FIGURE 2.1: Location of NHS Breast Screening Programme centres before 1998

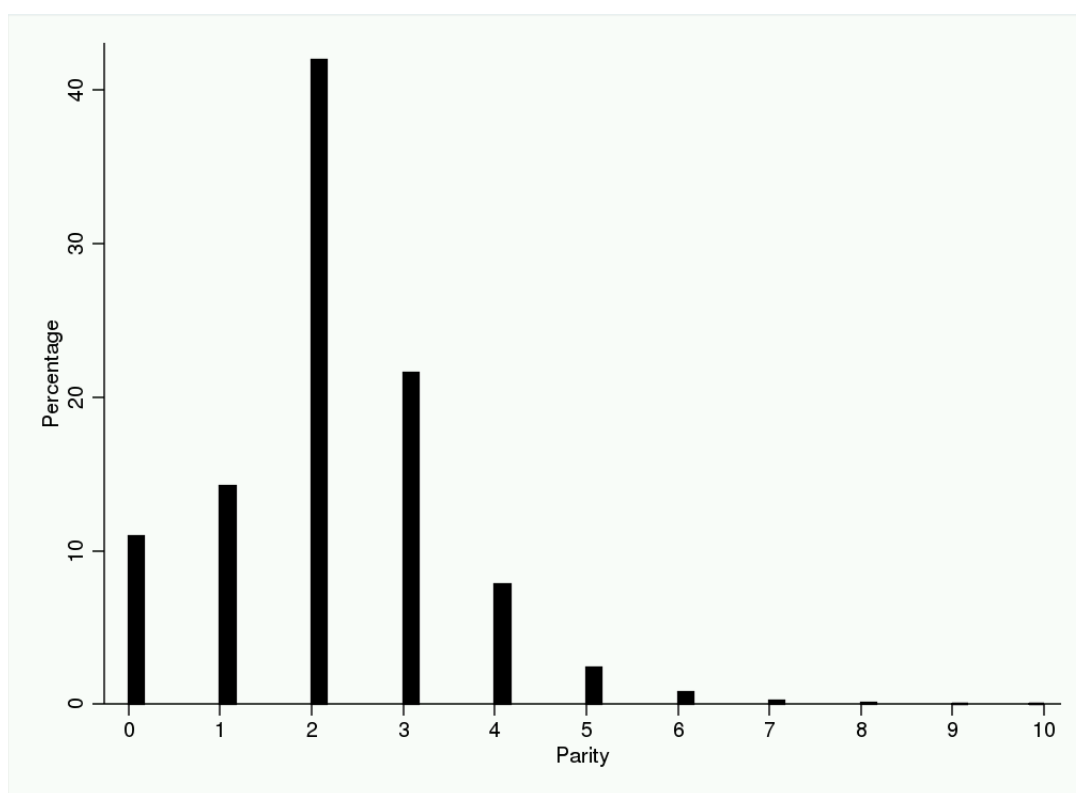


FIGURE 2.2: Histogram of parity among women in the Million Women Study

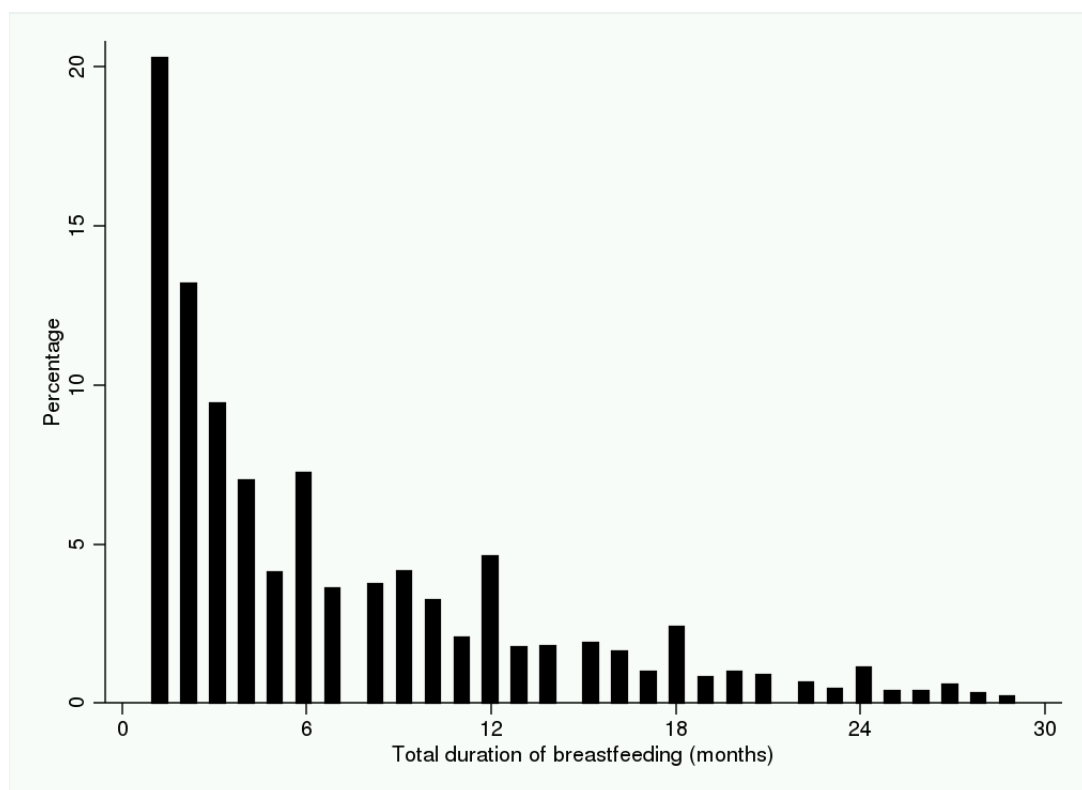


FIGURE 2.3: Distribution of total duration of breastfeeding among Million Women Study participants who breastfed

TABLE 2.1: Summary of published studies looking at reliability of long-term maternal recall of parity

Citation	Notes	n	Test-retest interval	How parity was defined	Sources of data	Measures of association
Tomeo (1999) Epidemiology 10:774 – [15]	Explored reproducibility of maternal recall of pregnancy-related events on two occasions two years apart. Mothers in the reproducibility study were on average 78 years at second survey, average recall interval was 57 years	146	2 years	On the 1993 questionnaire women were asked about pregnancies lasting 6 months or more, the same question as repeated on the 1995 questionnaire	Mothers of women enrolled in Nurses Health Studies asked to complete questionnaires on at least two separate occasions [2 years apart] (some women completed mail questionnaire, some had telephone interview)	Spearman correlation for "Previous live births (mean births +/- SD)" between recalls was 1.00
Buka (2000) Schizophhr Bull 26:335 – [16]	National Collaborative Perinatal Project (NCPP) Mothers of psychotic offspring and matched controls. All pregnancies selected for the validation study ended in a viable birth	56	25 years	Using a Pregnancy History Instrument Women were asked, "Before this child, how many times had you been pregnant?" "How many of these pregnancies ended with live births?", and "After this child, how many times had you been pregnant?" "How many of these pregnancies ended with live births?"	Medical records from NCPP Maternal recall interviews by telephone Results reported were event prevalence medical record and in maternal interview record, the sensitivity, specificity, and the kappa.	Among the patient sample the prevalence of parity = 0 or parity>4 was 25% (7 in each group) in both the chart and interview record, sensitivity and specificity were 100% and the kappa 1 Within the control sample the prevalence of parity = 0 was 18% (n=5), sensitivity & specificity were both 100%, and the kappa was equal to one. Among women with parity>4 one woman had conflicting data between her chart and interview record, as such the prevalence based on chart data was 18%, on interview data was 21% (n=6). The sensitivity was 100%, specificity 96%, and the kappa 0.89
Bosetti (2001) J Clin Epidemiol 54:902 – [17]	Reliability of self-report data on various reproductive factors and medical conditions. Controls from a hospital based case-control study of digestive tract neoplasm's contacted and re-interviewed at home	123	Between 12 and 24 months	Women were asked about "number of births and abortions"	171 women selected, 1243 interviewed twice about menstrual and reproductive histories	Of 123 women who were interviewed twice 89 had concordant responses giving a weighted percent of overall agreement of 99% and a kappa of 0.97

TABLE 2.2: Summary of published studies looking at reliability of long-term maternal recall of parity cont.

Citation	Notes	n	Test-retest interval	How parity was defined	Sources of data	Measures of association
Lin (2002) Annals of Epidemiology 12:242 - [18]	Test-retest reliability of reproductive and childhood social class variables Participants were 26 cases and 25 controls randomly selected for re-interview from a population-based case-control study of Hodgkin's disease	46	8 months	"Number of live births"	First interview in-person, second interview via telephone	Kappa or ICC reported for "the number of live births" using the full sample was 0.835 (n=45). Among cases it was 0.65 (n=24), and among controls it was 0.88 (n=21). These measures were not materially changed when stratified by age or education.
Langer (2003) Annals of Epidemiology 13:S107-S121. [19]	Test-retest reliability using a sub-set of women from Women's Health Initiative Observational Study who participated in Measurement Precision Study	564	Approximately 3 months (between 8 and 15 weeks)	"Ever pregnant" "Number of live births"	564 women completed self-administered baseline questionnaire as well as the repeat questionnaires	Kappa reported for "number of live births " 0.98
Slanger (2007) Annals of Epidemiology 17:993 - [20]	Test-retest reliability of women's self-reporting of reproductive factors in the setting of a large population-based case control study of HRT and breast cancer [MARIE]	123	Average time between interviews 40 weeks	"Number of pregnancies (0 to 7)"	Random sample of 480 women drawn from year 1 of MARIE, of which 1% selected for re-interview (138), final sample 123 women re-interviewed in-person	Agreement for number of pregnancies was 91%, kappa = 0.92

TABLE 2.3: Summary of published studies looking at reliability of long-term maternal recall of parity cont.

Citation	Notes	n	Test-retest interval	How parity was defined	Sources of data	Measures of association												
Ko (2008) Yebang Uihakhoe Chi 41:181 – [21]	Test-retest reliability of reproductive history questionnaire (KMCC)	386	8 years	“Number of children” Percent agreement for the number of children was calculated using the number of women, the percentage with the same number of children reported both times and the number of women derived from the reported error range	From the abstract it appears that information was gathered from interview – administered questionnaire	Percent agreement for pregnancy history was 99.7% with a kappa =0.67												
Lucas (2008) Journal of Clinical Epidemiology 61:945 - [22]	Reliability of self-reported reproductive factors among a cohort of postmenopausal women [EPIPorto study]	535	5 years (3 to 7 years)	“Number of live births”	In-person interview	Overall percent agreement for parity was 94.8%, kappa = 0.93(0.02) For individual parity levels it was, <table><tr><td>0</td><td>97.3 %</td></tr><tr><td>1</td><td>100.0 %</td></tr><tr><td>2</td><td>94.9 %</td></tr><tr><td>3</td><td>94.3 %</td></tr><tr><td>4</td><td>81.0 %</td></tr><tr><td>>4</td><td>84.4 %</td></tr></table>	0	97.3 %	1	100.0 %	2	94.9 %	3	94.3 %	4	81.0 %	>4	84.4 %
0	97.3 %																	
1	100.0 %																	
2	94.9 %																	
3	94.3 %																	
4	81.0 %																	
>4	84.4 %																	

TABLE 2.4: Summary of published findings on the validity of long-term maternal recall of breastfeeding

Citation	Notes	n	Recall interval	Questions on breastfeeding	Measures of association
Kark (1984) Journal of Epidemiology and Community Health 38:218 – [23]	“Validate a breast feeding history questionnaire applied to mothers about 20 years after the birth of their children against available routine records from mother and child clinics.”	74	20 years	“Did you breast feed your son/daughter when he/she was a baby?” “For how long did you breast feed him/her?”	Spearman correlation 0.86 (Pearson 0.82) Intercept near to the origin and the slope close to unity.
Vobecky (1988) J Clin Epidemiol 41:261 – [26]	Infant feeding information collected concurrently during first year post-partum and compared to data recalled eight years later. Self-administered questionnaire at monthly intervals until six months, thereafter at three monthly intervals. Second round of data collected via at-home interview using the same questions as in initial study	95	8 years	Not reported	Dispersion analysis and comparing means, also regression and correlation. Correlation coefficient 0.95 Slope 1.18 Intercept 0.82
Tienboon (1994) Aust J Nutr Dietet 51:25 – [25]	Comparison of maternal recall of infant feeding methods 15 years previously with information recorded contemporaneously on Infant Welfare Cards	144	14 years	Questionnaire including information on “the duration of breastfeeding (to the nearest month)”	Ever breastfed sensitivity 82% and specificity 93% (higher for first born children) Correlation for duration of breastfeeding 0.77
Promislow (2005) Am. J. Epidemiol. 161:289 – [24]	Women who completed prospective diary on breastfeeding duration during Menstruation and Reproductive History Study and also a follow up questionnaire on breastfeeding Used first matched breastfeeding event	140	34 to 50 years	“Did you breastfeed?” “If yes, for how many months?”	Sensitivity 94% Weighted Kappa Overall 0.55 (0.43, 0.67) By stratum of no. of births 1-2 0.33 (-0.00, 0.65) 3-4 0.53 (0.37, 0.69) 5-12 0.73 (0.16, 0.76)

TABLE 2.5: Summary of published findings on the reliability of long-term maternal recall of breastfeeding

Citation	Notes	n	Test-retest interval	Questions on breastfeeding	Measures of association
Rohan (1988) Journal of Clinical Epidemiology. 41:763 - [27]	Reliability study using 124 controls from two case-control studies of breast disease selected for telephonic re-interview.	107	8 to 17 months	Reported "Total duration of lactation" but actual questions not reported	Agreement 33 % exact agreement 58% agreed within 1 month Correlation coefficient 0.87
Tomeo (1999) Epidemiology 10:774 – [15]	Explored reproducibility of maternal recall of pregnancy-related events on two occasions two years apart. Validity study looked at recall data and collected data on pregnancy and delivery for one pregnancy.	146	2 years	Not reported	Spearman's Correlation Ever breastfed 0.89 Duration of breastfeeding 0.86
Langer (2003) Annals of Epidemiology 13:S107 - [19]	Test-retest reliability using a sub-set of women from Women's Health Initiative Observational Study who participated in Measurement Precision Study	564	3 months	" Number of months breastfed"	Kappa Months breastfed 0.89
Slanger (2007) Annals of Epidemiology 17:993-998. [20]	Test-retest reliability of women's self-reporting of reproductive factors in the setting of a large population-based case control study of HRT and breast cancer [MARIE] Duration of breastfeeding for first birth only	123	Average time between interviews 40 weeks	Reported ever breastfeeding and duration of breast-feeding of first child, actual questions not reported	Proportion overall agreement Duration of breastfeeding 90% Kappa Duration of breastfeeding 0.86
Ko (2008) Yebang Uihakhoe Chi 41:181 - [21]	Test-retest reliability of reproductive history questionnaire (KMCC) Article in Korean, abstract and results tables used	386	8 years	Article in Korean	Percentage agreement Breastfeeding history 96.1% Kappa Breastfeeding history 0.20

TABLE 2.6: Self-reported diabetes among Million Women Participants as reported at recruitment and at first repeat survey (about 3 years after recruitment)

Self-reported diabetes		At first repeat survey	
		"Have you had any of the following conditions [Diabetes] diagnosed for the first time in the last 5 years?"; "Are you now being treated for: [Diabetes] If yes, about how old were you when treatment began?"	
At recruitment		Yes	No
"Have you ever had Diabetes?"; "Are you now being treated for [Diabetes]?"	Yes	11,295 (83%)	2,317 (17%)
	No	8,037 (1.5%)	532,341 (98.5%)
Total		19,332	534,658
			13,612
			540,378
			553,990

Bibliography

- [1] The Million Women Study Collaborative Group. The Million Women Study: design and characteristics of the study population. *Breast Cancer Research: BCR*. 1999;1(1):73–80.
- [2] Blanks R, Moss S, Patnick J. Results from the UK NHS breast screening programme 1994–1999. *J Med Screen*. 2000 Dec;7(4):195–198.
- [3] NHS National Services Scotland. Information and Services Division, Scotland [online]; 2009. Internet [cited 02 Dec 2010]. Available from: http://www.isdscotland.org/isd/CCC_FirstPage.jsp.
- [4] NHS National Health Service. Health and Social Care Information Centre: Hospital Episode Statistics [online]; 2010. Internet [cited 02 Dec 2010]. Available from: <http://www.hesonline.nhs.uk>.
- [5] Kendrick S, Clarke J. The Scottish Record Linkage System. *Health Bulletin*. 1993 Mar;51(2):72–79.
- [6] World Health Organisation, editor. World Health Organization. International Statistical Classification of Diseases and Related Health Problems, vol. 1. Geneva: World Health OrganizationD; 1992.
- [7] Townsend P, Phillimore P, Beattie A. Health and deprivation. Routledge; 1988.
- [8] Rimm EB, Chan J, Stampfer MJ, Colditz GA, Willett WC. Prospective study of cigarette smoking, alcohol use, and the risk of diabetes in men. *BMJ*. 1995 Mar;310(6979):555–559.
- [9] Molarius A, Seidell JC, Kuulasmaa K, Dobson AJ, Sans S. Smoking and relative body weight: an international perspective from the WHO MONICA Project. *Journal of Epidemiology and Community Health*. 1997 Jun;51(3):252–260.
- [10] Doll R, Peto R, Boreham J, Sutherland I. Mortality in relation to smoking: 50 years' observations on male British doctors. *BMJ*. 2004 Jun;328(7455):1519.
- [11] Anand SS, Islam S, Rosengren A, Franzosi MG, Steyn K, Yusufali AH, et al. Risk factors for myocardial infarction in women and men: insights from the INTERHEART study. *European Heart Journal*. 2008 Apr;29(7):932–940.
- [12] Fletcher C, Peto R, Speizer F, Tinker C. A follow-up study of the natural history of obstructive bronchitis. Oxford University Press; 1976.

- [13] MacMahon S, Peto R, Collins R, Godwin J, MacMahon S, Cutler J, et al. Blood pressure, stroke, and coronary heart disease : Part 1, prolonged differences in blood pressure: prospective observational studies corrected for the regression dilution bias. *The Lancet*. 1990 Mar;335(8692):765–774.
- [14] Banks E, Richardson A, Beral V, Crossley B, Simmonds M, Hilton E, et al. Effect on attendance at breast cancer screening of adding a self administered questionnaire to the usual invitation to breast screening in southern England. *Journal of Epidemiology and Community Health*. 1998 Feb;52(2):116–119.
- [15] Banks E, Beral V, Cameron R, Hogg A, Langley N, Barnes I, et al. Comparison of various characteristics of women who do and do not attend for breast cancer screening. *Breast Cancer Research: BCR*. 2002;4(1):R1.
- [16] Bowling A. In: *Research methods in health : investigating health and health services*; 2009. .
- [17] Hennekens CH, Buring JE, Mayrent SL. *Epidemiology in medicine*. 1st ed. Philadelphia: Lippincott Williams & Wilkins; 1987.
- [18] Tomeo CA, Rich-Edwards JW, Michels KB, Berkey CS, Hunter DJ, Frazier AL, et al. Reproducibility and Validity of Maternal Recall of Pregnancy-Related Events. *Epidemiology*. 1999 Nov;10(6):774–777.
- [19] Buka SL, Goldstein JM, Seidman LJ, Tsuang MT. Maternal Recall of Pregnancy History: Accuracy and Bias in Schizophrenia Research. *Schizophr Bull*. 2000 Jan;26(2):335–350.
- [20] Bosetti C, Tavani A, Negri E, Trichopoulos D, Vecchia CL. Reliability of data on medical conditions, menstrual and reproductive history provided by hospital controls. *Journal of Clinical Epidemiology*. 2001 Sep;54(9):902–906.
- [21] Lin SS, Glaser SL, Stewart SL. Reliability of Self-Reported Reproductive Factors and Childhood Social Class Indicators in a Case-Control Study in Women. *Annals of Epidemiology*. 2002 May;12(4):242–247.
- [22] Langer RD, White E, Lewis CE, Kotchen JM, Hendrix SL, Trevisan M. The women's health initiative observational study: baseline characteristics of participants and reliability of baseline measures. *Annals of Epidemiology*. 2003 Oct;13(9, Supplement 1):S107–S121.
- [23] Slanger T, Mutschelknauss E, Kropp S, Braendle W, Flesch-Janys D, Chang-Claude J. Test-Retest Reliability of Self-Reported Reproductive and Lifestyle Data in the Context of a German Case-Control Study on Breast Cancer and Postmenopausal Hormone Therapy. *Annals of Epidemiology*. 2007 Dec;17(12):993–998.
- [24] Ko K, Park S, Kim Y, Bae J, Jun J, Gwack J, et al. Reliability of a questionnaire for women's reproductive history. *Journal of Preventive Medicine and Public Health Korean*. 2008;41(3).
- [25] Lucas R, Azevedo A, Barros H. Self-reported data on reproductive variables were reliable among postmenopausal women. *Journal of Clinical Epidemiology*. 2008 Sep;61(9):945–950.
- [26] Kark JD, Troya G, Friedlander Y, Slater PE, Stein Y. Validity of maternal reporting of breast feeding history and the association with blood lipids in 17 year olds in Jerusalem. *Journal of Epidemiology and Community Health*. 1984;38(3):218 –225.

- [27] Promislow JHE, Gladen BC, Sandler DP. Maternal Recall of Breastfeeding Duration by Elderly Women. *Am J Epidemiol*. 2005 Feb;161(3):289–296.
- [28] Tienboon P, Rutishauser I, Wahlqvist M. Maternal recall of infant feeding practices after an interval of 14 to 15 years. *Aust J Nutr Dietet*. 1994;51:25–27.
- [29] Vobecky JS, Vobecky J, Froda S. The reliability of the maternal memory in a retrospective assessment of nutritional status. *Journal of Clinical Epidemiology*. 1988;41(3):261–265.
- [30] Rohan TE, Record SJ, Cook MG. Repeatability of interview-derived socio-demographic and medical information. *Journal of Clinical Epidemiology*. 1988;41(8):763–770.
- [31] Feinstein AR, Cicchetti DV. High agreement but low kappa: I. The problems of two paradoxes. *Journal of Clinical Epidemiology*. 1990;43(6):543–549.
- [32] Witts LJ. *Medical Surveys and Clinical Trials: Some Methods And*. 2nd ed. Oxford medical publications. Oxford University Press: London; 1964.
- [33] Liddell F. The development of cohort studies in epidemiology: A review. *Journal of Clinical Epidemiology*. 1988;41(12):1217–1237.
- [34] Reeves GK, Pirie K, Beral V, Green J, Spencer E, Bull D, et al. Cancer incidence and mortality in relation to body mass index in the Million Women Study: cohort study. *BMJ*. 2007 Dec;335(7630):1134.
- [35] Altman DG, Bland JM. Measurement in Medicine: The Analysis of Method Comparison Studies. *Journal of the Royal Statistical Society Series D (The Statistician)*. 1983 Sep;32(3):307–317.
- [36] Bland JM, Altman DG. Comparing methods of measurement: why plotting difference against standard method is misleading. *The Lancet*. 1995 Oct;346(8982):1085–1087.
- [37] Heliövaara M, Aromaa A, Klaukka T, Knekt P, Joukamaa M, Impivaara O. Reliability and validity of interview data on chronic diseases The mini-Finland health survey. *Journal of Clinical Epidemiology*. 1993 Feb;46(2):181–191.
- [38] Bowlin SJ, Morrill BD, Nafziger AN, Lewis C, Pearson TA. Reliability and changes in validity of self-reported cardiovascular disease risk factors using dual response: The behavioral risk factor survey. *Journal of Clinical Epidemiology*. 1996 May;49(5):511–517.
- [39] The Italian Longitudinal Study on Aging Working Group. Prevalence of chronic diseases in older Italians: comparing self-reported and clinical diagnoses. *International Journal of Epidemiology*. 1997 Oct;26(5):995–1002.
- [40] Goldman N, Lin I, Weinstein M, Lin Y. Evaluating the quality of self-reports of hypertension and diabetes. *Journal of Clinical Epidemiology*. 2003 Feb;56(2):148–154.
- [41] Okura Y, Urban LH, Mahoney DW, Jacobsen SJ, Rodeheffer RJ. Agreement between self-report questionnaires and medical record data was substantial for diabetes, hypertension, myocardial infarction and stroke but not for heart failure. *Journal of Clinical Epidemiology*. 2004 Oct;57(10):1096–1103.

- [42] Simpson CF, Boyd CM, Carlson MC, Griswold ME, Guralnik JM, Fried LP. Agreement Between Self-Report of Disease Diagnoses and Medical Record Validation in Disabled Older Women: Factors That Modify Agreement. *Journal of the American Geriatrics Society*. 2004;52(1):123–127.
- [43] Banks E, Beral V, Cameron R, Hogg A, Langley N, Barnes I, et al. Agreement between general practice prescription data and self-reported use of hormone replacement therapy and treatment for various illnesses. *Journal of Epidemiology and Biostatistics*. 2001;6(4):357.
- [44] Bergmann MM, Jacobs EJ, Hoffmann K, Boeing H. Agreement of Self-Reported Medical History: Comparison of an In-Person Interview with a Self-Administered Questionnaire. *European Journal of Epidemiology*. 2003;19(5):411–416.
- [45] Walker M, Shaper AG, Wannamethee SG, Whincup PH. Measuring the prevalence of disease in middle-aged British men. *Journal of the Royal College of Physicians of London*. 1999 Aug;33(4):351–358.
- [46] ONS Office for National Statistics. Mortality statistics:General Statistics [online]; 2010. Internet [cited 02 Dec 2010]. Available from: <http://www.statistics.gov.uk/statbase/Product.asp?vlnk=618&More=Y>.
- [47] GRO General Register Office for Scotland. Scottish Mortality Statistics [online]; 2010. Internet [cited 02 Dec 2010]. Available from: <http://www.gro-scotland.gov.uk/statistics/deaths/index.html>.
- [48] Silver L, Heneghan C, Mehta Z, Banning A, Rothwell P. Substantial underestimation of incidence of acute myocardial infarction by hospital discharge diagnostic coding data: a prospective population-based study. *Heart (British Cardiac Society)*. 2009;95(2).
- [49] Patrick DL, Cheadle A, Thompson DC, Diehr P, Koepsell T, Kinne S. The validity of self-reported smoking: a review and meta-analysis. *Am J Public Health*. 1994 Jul;84(7):1086–1093.
- [50] Steffensen FH, Lauritzen T, Sørensen HT. Validity of self-reported smoking habits [letter]. *Scandinavian Journal of Primary Health Care*. 1995 Jul;(13).
- [51] Brigham J, Lessov-Schlaggar CN, Javitz HS, Krasnow RE, Tildesley E, Andrews J, et al. Validity of Recall of Tobacco Use in Two Prospective Cohorts. *American Journal of Epidemiology*. 2010 Oct;172(7):828–835.
- [52] Wareham NJ, Jakes RW, Rennie KL, Schuit J, Mitchell J, Hennings S, et al. Validity and Repeatability of a Simple Index Derived from the Short Physical Activity Questionnaire Used in the European Prospective Investigation into Cancer and Nutrition (EPIC) Study. *Public Health Nutrition*. 2003;6(04):407–413.
- [53] Armstrong MEG, Cairns B, Beral V. Reliability of self-reported physical activity over time and estimated energy expenditure in the Million Women Study. Unpublished. 2010;.
- [54] Bohlscheid-Thomas S, Hoting I, Boeing H, Wahrendorf J. Reproducibility and relative validity of food group intake in a food frequency questionnaire developed for the German part of the EPIC project. *European Prospective Investigation into Cancer and Nutrition. International Journal of Epidemiology*. 1997;26 Suppl 1:S59–70.

- [55] Ocke M. The Dutch EPIC food frequency questionnaire. II. Relative validity and reproducibility for nutrients. *International Journal of Epidemiology*. 1997;26(90001):49S–58.
- [56] Roddam AW, Spencer E, Banks E, Beral V, Reeves G, Appleby P, et al. Reproducibility of a Short Semi-Quantitative Food Group Questionnaire and Its Performance in Estimating Nutrient Intake Compared with a 7-Day Diet Diary in the Million Women Study. *Public Health Nutrition*. 2005;8(02):201–213.
- [57] Easton DF, Peto J, Babiker AGAG. Floating absolute risk: An alternative to relative risk in survival and case-control analysis avoiding an arbitrary reference group. *Statistics in Medicine*. 1991 Jul;10(7):1025–1035.

Chapter 3

Trends in breastfeeding in the United Kingdom, comparing breastfeeding reported in the Million Women Study with historical data

Abstract

Objectives: *To describe historical trends in the annual proportion of women who breastfed in the UK during the second half of the 20th century using data from the Million Women Study, and to compare these to estimates from other data sources*

Design & setting: *Data from the Million Women Study, and the 1958 and 1970 birth cohorts, and survey data from the National Infant Feeding surveys in 1975, 1980, and 1985.*

Main outcome measures: *Proportion of women who reported breastfeeding a child born in the study year of interest, and the duration of breastfeeding for that child*

Results: *The average recall interval for breastfeeding was 30 years for women participating in the Million Women Study, 6 years for women who participated in one of the birth cohort studies, and less than one year for women who participated in one of the National Infant Feeding surveys. For calendar year examined, the proportion of mothers from the Million Women Study who reported breastfeeding was remarkably similar to women from the other studies. For example, 36% of Million Women Study participants with a child born in 1970 reported that they breastfed that child compared to 37% for women participating in the 1970 birth cohort study. Taken together the data show that breastfeeding rates fell from a high of 74% for births in 1955 to a low of 36% for births in 1970, and then increased again to nearly 70% for births in 1980. The trend in breastfeeding rates over time was similar across socioeconomic groups, however women from the lowest socioeconomic groups had the lowest breastfed rates. There was some evidence that the difference in breastfeeding rates between socioeconomic groups increased between the 1950s (6%) and the 1980s (23%). It was also noted that women were more likely to have ever breastfed first born than later born children.*

Conclusions: *Breastfeeding in the United Kingdom declined for births from 1955 until 1970 when only about one third of women reported breastfeeding their children. Thereafter the rates steadily increased until by the mid-1980s more than two thirds of women reported breastfeeding their children. Data from the Million Women Study (long-term maternal recall of breastfeeding) are highly comparable to data from the 1958 and 1970 birth cohorts, and the early National Infant Nutrition Feeding surveys*

3.1 Introduction

Trends in the proportion of women breastfeeding changed dramatically in developed-economy countries during the 20th century. [1] Evidence suggesting significant variation in breastfeeding initiation rates and duration usually come from time point data which were sparse in the United Kingdom (UK) before the National Infant Feeding Surveys started in 1975. [2–4] The Million Women Study contains sufficient data to look at breastfeeding estimates and duration by year of birth of the child for a period spanning 40 years (from 1950 to 1990). Population-based studies conducted in the same time period which contain information on childbearing and breastfeeding are the 1958 Birth Cohort, [2, 5] the 1970 British Birth Cohort, [6] and the National Infant Feeding surveys from 1975 to 1985. [7–9] To my knowledge, no previous study has described historical trends in the annual proportions of women who breastfed over such a long period, nor compared estimates across several independent sources of data.

3.2 Methods

3.2.1 Study Population

Million Women Study

The Million Women Study has been described in Chapter 2 (page 18). Briefly, it is a population-based cohort of middle-aged women recruited via the National Health Service (NHS) Breast Screening Programme in England and Scotland between 1996 and 2001. A total of 1.3 million women completed the first study questionnaire which asked about their medical and reproductive histories, including the number children

they had given birth to (including still births) and for each birth, whether or not they had breastfed (and if so, the duration of breastfeeding in months). Women were also asked about other personal, sociodemographic, and lifestyle characteristics. It can be seen from figure 3.1 (page 79) that parous participants mostly had births between the mid-1950s and the mid-1980s (< 1% of women had a birth before 1950 or after 1985). It was therefore relevant to look for other data sources of infant feeding around these years i.e. 1958, 1970, 1975, 1980, 1985.

Birth cohorts

The 1958 and 1970 birth cohorts each enrolled all children born in England, Wales and Scotland during a one week period in March 1958 and April 1970 respectively. [5, 6] For both cohorts information on parental socio-demographic, reproductive and pregnancy histories was collected via a midwife interview with the mother at birth¹. For the 1958 birth cohort² follow-up interviews of the parents and the index child occurred in 1965, 1969, 1975, and 1991; [5] mothers were asked if they breastfed the index child and if so, for how long in 1965 and again in 1991. For the 1970 birth cohort³ follow-up interviews of the parents occurred in 1975, 1980, 1986, 1996, 2000, and 2004; mothers were asked if they breastfed the index child, and if so for how long in 1975, and again in 1986. [6]

National Infant Feeding Surveys

National Infant Feeding Surveys began in England and Wales in 1975 in response to a request by the Department of Health and Social Security for evidence and best practice

¹for the 1970 birth cohort at recruitment women were also asked about attempted lactation and infant feeding in the first seven days postpartum

²Previously known as the National Child Development Study

³previously known as the British Cohort Study or British Births Survey

in regard to infant nutrition. [10] The surveys are conducted every 5 years. For this analysis I used published data from the 1975, 1980, and 1985 surveys. [7–9] Survey researchers used birth registries to generate population representative samples. Mothers were contacted and asked to complete survey questionnaires via post at six weeks, four months, and nine months post-partum. Each questionnaire asked mothers to report their current infant feeding practices. Women who initially reported they were breastfeeding the index child but who stopped before the next survey wave were asked to recall for how long they had breastfed the index child. The questionnaires also asked about mothers experience of breastfeeding any of their other children. Information on sociodemographic and pregnancy related factors was also collected.

3.2.2 Data Definitions

Table 3.1 on page 80 summarises the breastfeeding questions asked in the Million Women Study, the 1958 and 1970 birth cohort studies, and the infant feeding surveys.

Million Women Study

As can be seen in table 3.1 (page 80) women were asked at recruitment to report for each child, the child's date of birth, whether or not the child was breastfed and, if yes, the duration of breastfeeding. Using each child's date of birth, I created groups of children born for each year (from the 1st of January to the 31st December inclusive) from 1950 until 1985. I used mother's report of whether or not a child born in a given year was breastfed to create a binary "ever breastfed" variable for that child. For each child breastfed, women were instructed to fill in "1" for any duration of breastfeeding

of 1 month or less and for any duration of breastfeeding longer than a month an open space was provided to fill in how long that child was breastfed in months.

Birth cohort studies

In both of the birth cohort studies women were asked to recall how they fed the index child (table 3.1, page 80). Women were asked if they breastfed the child and, if so, for how long. For the 1958 and 1970 birth cohorts, data were extracted from the 1965 and 1975 follow-up surveys respectively.

National Infant Feeding Surveys

Women with a child born in the year of the survey were asked about how they fed their infant at six weeks, four months, and nine months postpartum. The format and wording of the questions on breastfeeding was broadly the same for each survey (table 3.1, page 80). [11–13] I extracted data from tables, figures and text found in published surveys for 1975, 1980, and 1985. [11–13]

Comparing durations of breastfeeding across data sources

There were no standard categories for the duration of breastfeeding that that could be applied across all studies due to the use of different measures of the duration for which women breastfed the index child. I created standard measures of the duration of breastfeeding in the Million Women Study using the definition appropriate for each reference study. For example, for the 1958 birth cohort women were asked, if they breastfed the index child, to categorise the duration they breastfed as either less than one month or

more than one month. For Million Women Study participants who reported breastfeeding a child born in 1958 I created similar categories for duration of breastfeeding (one month or less, or more than one month). For the 1970 birth cohort, women were asked, if they breastfed the index child, to categorise the duration they breastfed as less than 1 month, 1 month or more but less than 3 months, or 3 months or more. For the Million Women Study participants who reported breastfeeding a child born in 1970, I created the following categories for breastfeeding duration, breastfed for one month or less, for more than one month but less than three months, or for more than three months.

3.2.3 Statistical Analysis

I restricted the analysis period for the Million Women Study data to between 1950 and 1985 inclusive as most women gave birth during this time. For ease of comparison between datasets I calculated the proportion of women who breastfed and the duration of breastfeeding for the years 1950, 1955, 1958, 1960, 1965, 1970, 1975, 1980, and 1985. To calculate estimates of the proportion of women who breastfed in a given year, I used the number of children breastfed in that year divided by the total number of children born in that year. To calculate estimates for breastfeeding duration I created broad categories of duration similar to the appropriate reference study and used the total number of children as the denominator. I presented the estimates of the proportion of women who breastfed in a given year separately by birth order (first born or later born) and by socioeconomic group (in tertiles based on the Townsend deprivation index). [14]

For the 1958 and 1970 birth cohorts I accessed the datasets from the Centre for Longitudinal Studies available through the UK Data Archive. [15, 16] Then I analysed the data in STATA version 11 (Stata corporation, TX, USA) and calculated the proportion

of women who breastfed a child born in these years. My estimates agreed with those from other published data that were based on the same datasets. [17] For the 1975, 1980, and 1985 National Infant Nutrition Surveys I used the estimates as published in the survey reports. [7–9] I compared the consistency of reported estimates for these years over time using survey reports from 1990 and 1995. [18, 19]

I compared estimates of the proportion of women who breastfed calculated from the Million Women Study data for the years 1958 and 1970 with data from the 1958 and 1970 birth cohorts. To compare estimates of breastfeeding duration I used similarly defined categories between the Million Women Study and the appropriate reference study. I also compared study estimates by birth order. I compared estimates of the proportion of women who breastfed calculated from the Million Women Study data for 1975, 1980, and 1985 with the national infant feeding surveys for those years, and again compared estimates by birth order. It was not possible for me to directly compare estimates for breastfeeding duration due to differences in study methodologies.

3.3 Results

Characteristics of the study populations

Table 3.3 on page 81 shows selected characteristics of Million Women Study participants for selected years of interest. The number of participants contributing data in a given year ranged from 1,720 for women who gave birth in 1950 to 140,404 for women who gave birth in 1965. As calendar time progressed the number of women reporting a birth increased, peaking between 1960 and 1975. Correspondingly the proportion of women with a first birth decreased, and the average age of the mother and

her age at first birth (if she was primiparous) also increased. Relatively few Million Women Study participants reported a birth in 1950 or in 1985 (<1%). Almost three-quarters of women (73%) reporting a birth in 1950 were first time mothers, and these women were younger than the study population average (20 years [SD 4] compared to 24 years [SD 4]. Very few women who reported a birth in 1980 reported that this was their first child (17%), and these women were older than the study average (29 years [SD 7]).

Table 3.2 on page 81 shows selected study and participant characteristics for each of the birth cohorts and the infant feeding surveys. The 1958 birth cohort included 15,425 women, the 1970 birth cohort included 13,071. The mean age of women who participated in the birth cohort studies was about 26 years. Slightly more than one third (37%) of women participating in the birth cohort studies were first time mothers. The breastfeeding recall interval was on average 7 years for women who participated in the 1958 birth cohort, 5 years for the 1970 birth cohort, and 40 years [SD 1] for women from the Million Women Study. The National Infant Feeding surveys used in this analysis sampled 1544, 3755, and 4671 women respectively. One third of women included in each of these surveys were aged 20 to 24 years, and slightly less than half were first time mothers. Information on breastfeeding was based on recent maternal recall (less than one year).

It can be seen from tables 3.3 and 3.2 (pages 81 and 81 respectively) that the characteristics of Million Women Study participants for the years of interest and the characteristics of mothers who participated in either the birth cohorts or the infant feeding surveys were similar.

Comparing estimates across studies

Breastfeeding incidence

Table 3.4 (page 82) shows the estimates of the proportion of women who breastfed a child born in a given year for the Million Women Study, the birth cohort studies, and the National Infant Feeding surveys. The proportion of women who breastfed a child born in a given year were similar across comparable studies. The proportion of women who breastfed a child born in a given year by birth order were also similar across comparable studies (table 3.4, page 82).

Women with a child born in 1958 Million Women Study participants who reported a birth in 1958 were on average younger and more likely to be first time mothers than women who participated in the 1958 birth cohort. The proportions of UK women with a birth in 1958 who reported breastfeeding a child born in that year were comparable for these two data sources (70% according to 1950 birth cohort data compared to 69% using data from the Million Women Study).

Women with a child born in 1970 For women with a child born in 1970 the average age of first birth and the proportions of women who were first time mothers were similar for women from the Million Women Study and for women who participated in the 1970 birth cohort (26 years, 37%). Again, the proportions of UK women with a birth in 1970 who reported breastfeeding a child born in that year were similar for the two data sources (38% of women who participated in the 1970 birth cohort, 36% of women from the Million Women Study).

Women with a child born in 1975, 1980, or 1985 Million Women Study participants who reported having a birth in 1975, 1980, or 1985 were older and less likely to be first time mothers than mothers who participated in a corresponding national infant feeding survey. The proportions of UK women who reported breastfeeding a child born in 1975, 1980, or 1985 were similar for data from the Million Women Study and each of the infant feeding surveys. Data from both data sources suggest that the proportion of UK women who reported breastfeeding increased between 1975 and 1985.

Breastfeeding duration

For women who breastfed a child born in one of the years of interest, table 3.5 (page 82) shows the average duration of breastfeeding. Estimates calculated using Million Women Study data are slightly lower than the other studies for the proportion of women who breastfed in a given year, and higher than the other studies for longer durations of breastfeeding.

Due to differences in methodologies I was not able to directly compare the duration of breastfeeding between the Million Women Study and the infant feeding surveys.

Trends in breastfeeding in the UK over time

Figure 3.2 (page 83) shows the proportion of UK women who reported breastfeeding a child born in a specific year for each study considered in this analysis. For the Million Women Study, data are presented for the years 1950 to 1985 at 5 year intervals, and also for 1958. Data from the Million Women Study agrees well with data from studies at comparable time points. As can be seen from the figure the proportions of women who reported breastfeeding a child born in a given year declined from more than 70%

in the early 1950s to just over one third in the early 1970s. Thereafter the trend reversed with the proportions of women reporting breastfeeding increasing for children born from the mid-nineteen seventies onwards. By 1985 the proportion of women reporting breastfeeding a child born in that year was again similar to, though still lower than, the proportion of women who reported breastfeeding a child in the 1950s.

Figure 3.3 (page 83) shows the data from the Million Women Study by socioeconomic tertile. The changes in breastfeeding over time were similar within each socioeconomic group. Of note the difference in the proportion of women reporting breastfeeding a child in a given year between women in the highest socioeconomic tertile and the lowest socioeconomic tertile appeared to increase over calendar time. Figure 3.4 (page 84) shows this difference. These data suggest that until the nineteen seventies the difference between the proportion of women who breastfed between the highest socioeconomic and lowest socioeconomic groups was reasonably stable at around 10%. Thereafter, the difference between these groups increased to about 17% by the mid-nineteen eighties. Data from the birth cohorts and the infant feeding surveys (although not directly comparable) support the idea that the difference in the proportion of women who breastfed between higher and lower socioeconomic groups increased over time. The difference in the proportion of women who breastfed between women from non-manual social classes and manual social classes (as defined by the social class of her husband/partner) was 6% among participants of the 1958 birth cohort, and among participants of the 1970 birth cohort it was 13%. Data from the National Infant Feeding surveys suggest that by 1980 the difference was 23%, thereafter remaining reasonably stable at just over 20%.

3.4 Discussion

Summary

Breastfeeding data from the Million Women Study were highly comparable to data from the 1958 and 1970 birth cohorts, and the early National Infant Nutrition Feeding surveys. All of the data suggest that breastfeeding trends in the United Kingdom have undergone a remarkable shift over the last half of the twentieth century. Specifically, the proportion of women breastfeeding fell between the 1950s and the 1970s (at its lowest point only one third of women reported breastfeeding) before starting to increase. By the mid 1980s the proportion of women who reported breastfeeding was nearly as high as the 1950s. There is some evidence to suggest that while the overall proportion of women who breastfed increased from the 1970s onwards the distribution by socioeconomic group became more unequal. In other words, over time many more women from higher socioeconomic groups reported breastfeeding their children as compared to women from lower socioeconomic groups.

Findings in relation to other studies

To my knowledge there is no published literature on breastfeeding trends in the United Kingdom before the National Infant Feeding surveys were started in 1975 (prior to this estimates were based on birth cohort and similar data). [10] There is parallel evidence from countries similar to the UK including the USA, Australia and New Zealand that shows breastfeeding rates fell in the middle of the 20th century, probably due to perceptions of best infant feeding practices that were endorsed, if not actively encouraged, by

the medical profession and published in the popular media. [20–22] Similar to my findings, these countries all experienced a rebound in breastfeeding rates from the 1980s onwards. [1, 23]

Strengths and weaknesses

Although the data from all of the studies used in this analysis are from self-report and may be affected by systematic bias we have already seen that women's recall of breastfeeding generally has good *reliability* and *validity* (see section 2.4.2.2, page 34 in chapter 2), and the consistency of the data from multiple independent sources argue against substantial systematic bias.

These findings may be limited by the nature of the different sources and types of data. For example, the the birth cohorts and feeding surveys were designed to answer questions about infants rather than mothers. In addition I was limited in my ability to compare duration of breastfeeding across studies because of the different ways in which duration had been assessed.

Strengths of this analysis include the use of multiple sources of data which were independent in time and geography, the size of the Million Women Study and detailed information on breastfeeding spanning a period of almost 40 years. In addition the 1958 and 1970 birth cohorts asked about breastfeeding in several waves of follow up and investigators found similar estimates at multiple time points.

Conclusions

Data from multiple sources suggest that breastfeeding rates in the United Kingdom changed substantially during the 20th century. More than two thirds of women reported breastfeeding in the mid- to late 1950s. By 1970 the opposite was true, only a third of women reported breastfeeding. From the 1980s onwards and possibly as a result of international and national breastfeeding promotion campaigns rates of breastfeeding recovered so that by the early 1990s two thirds of women reported breastfeeding. In addition to changes in the overall pattern of breastfeeding that were seen in all socioeconomic groups there was some evidence that inequality in breastfeeding rates between the lowest and highest socioeconomic groups increased over time. Longterm recalled breastfeeding data from the Million Women Study are comparable to short and medium term recalled breastfeeding data from other nationally representative UK data.

3.5 Chapter summary

In this chapter I explored how breastfeeding data from the Million Women Study compared to other available national data on breastfeeding. I found that for the years compared, breastfeeding data from the Million Women Study compared very well to data from the 1958 and 1970 birth cohorts, and the 1975, 1980, and 1985 National Infant Feeding surveys. I also found that trends in breastfeeding in the UK changed over the course of the latter half of the 20th century. Socioeconomic status was an important co-factor, and I found evidence that the impact of socioeconomic status on the proportion of women of breastfed and breastfeeding duration may vary over calendar time. In general, most Million Women Study participants breastfed for short durations thus for chapters analysing the associations between breastfeeding and BMI, diabetes, or ischaemic heart disease will have substantial power to explore the effects of short durations of breastfeeding on risk. The results associated with longer durations of breastfeeding will need to be cautiously interpreted. As shown in this chapter breastfeeding patterns are susceptible to change, as such the relevance of breastfeeding as a potentially modifiable risk factor for various diseases that could be mobilised as a public health intervention will likely depend on the prevailing breastfeeding rates.

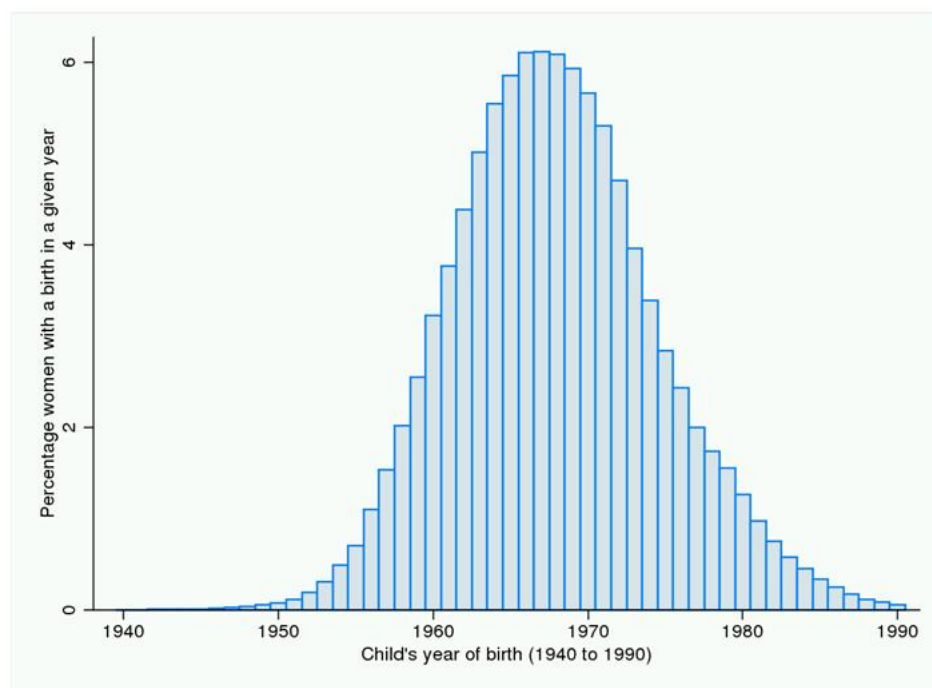


FIGURE 3.1: Histogram of calendar years in which parous Million Women Study participants reported having a birth

TABLE 3.1: Breastfeeding questions from the Million Women Study recruitment questionnaire (aqua), the 1958 and 1970 birth cohort questionnaires, and the 1975, 1980, and 1985 National Infant Feeding survey questionnaires

Study	Child's year of birth	Breastfeeding asked	Questions on breastfeeding
Birth cohort	1958	<i>in 1965</i>	"Was the child breast fed (partly or wholly) as a baby?" [No] [Yes - under 1 month] [Yes - over one month] [Don't know]
	1970	<i>in 1975</i>	" Was N breast fed partly or wholly, even for a few days?" [Yes - for less than one month] [Yes - for 1 month or more but less than 3 months] [Yes - for 3 months or more] [Yes - but can't remember how long] [No, was not breastfed at all] [Not known]
Infant feeding surveys	1975, 1980, 1985	<i>at 6 weeks</i>	" Did you ever put your baby to the breast?" [Yes (even if it was once only)] [No, never] " At the moment is your baby....." [bottle fed] [breast fed] [or both?]
			"How old was your baby when you last breastfed him/her?" [less than 1 week] [1 week but less than 2 weeks] [2 weeks but less than 6 weeks] [6 weeks but less than two months] [2 months or more]
		<i>at 4 months</i>	"Are you still breastfeeding him/her at all?" [Yes] [No]
			" How old was your baby when you last breast fed him/her?" [less than 2 months old] [2 months, but less than 3 months] [3 months, but less than 4 months] [4 months or more]
		<i>at 9 months</i>	"Are you still breastfeeding him/her at all?" [Yes] [No]
Million Women Study	years 1950 to 1985	<i>at recruitment</i>	"How old was your baby when you last breast fed him/her?" [less than 6 months old] [6 months, but less than 9 months] [9 months or more]
			" When was each child born, and for how many months did you breastfeed each child, if at all" [Months that you breastfed each child; put "0" if you did not breastfeed that child and "1" if you breastfed for a month or less"

TABLE 3.2: Characteristics of studies used in this analysis - 1958 and 1970 Birth Cohorts and the 1975, 1980, and 1985 National Infant Feeding surveys

Study	National birth cohort		National Infant Feeding Surveys		
Study Design	Prospective cohort	Prospective cohort	Population-based survey	Population-based survey	Population-based survey
Data sources	Interview & questionnaire	Medical record & questionnaire	Postal survey	Postal survey	Postal survey
Breastfeeding information	Maternal recall	Maternal recall	Contemporary practice	Contemporary practice	Contemporary practice
Age at entry to study * [SD]	27 [6]	26 [7]	32% (20 – 24 yrs)	31% (20 – 24 yrs)	30% (20 – 24 yrs)
Year (s) of child's birth	1958	1970	1975	1980	1985
Year of survey	1965	1975	1975	1980	1985
Response (%)	86	79	Overall response between 89-93%		
Number of participants	15,425	13,071	1,544	3,755	4,671
Study child is first born (%)	37	37	46	44	46
Recall interval in years (Mean [SD])	7	5	< 1	< 1	< 1

TABLE 3.3: Characteristics of studies used in this analysis - the Million Women Study

Study	Million Women Study								
Study Design	Prospective cohort								
Data sources	Self-administered questionnaire								
Breastfeeding information	Maternal recall								
Age at entry to study [SD]	56 [5]								
Year (s) of child's birth	1950	1955	1958	1960	1965	1970	1975	1980	1985
Year of survey	Between 1996 to 2001								
Number of participants	1720	15,894	47,033	75,564	140,404	140,274	71,752	32,186	8497
Response (%)	71% of women who attended for screening								
Study child is first born (%)	73	73	71	56	45	37	30	21	17
Age at first birth (Mean [SD])	21 [3]	20 [2]	21 [2]	21 [3]	22 [3]	23 [3]	25 [4]	27 [5]	29 [7]
Recall interval in years (Mean [SD])	48 [1]	43 [1]	40 [1]	38 [1]	33 [1]	29 [1]	24 [1]	19 [1]	14 [2]

TABLE 3.4: Comparison of estimates of the proportion of women breastfeeding, and of the estimates by birth order for comparable years between the Million Women Study and the 1958, and 1970 birth cohorts, and the 1975, 1980, and 1985 National Infant Feeding surveys

Proportion of women who reported breastfeeding a child born in the study year			
All women			
Year of interest	Million Women Study	Birth cohort	National infant feeding survey
1958	69	70	
1970	36	38	
1975	48		51
1980	65		67
1985	66		65
If child is first born			
Year of interest	Million Women Study	Birth cohort	National infant feeding survey
1958	72	73	
1970	43	45	
1975	60		-
1980	74		74
1985	73		69

TABLE 3.5: For women who breastfed, comparison of categories of breastfeeding duration for comparable years between the Million Women Study and the 1958, and 1970 birth cohorts

Duration of breastfeeding			
Year of interest	Breastfeeding duration	Million Women Study	Birth cohort
1958	> 1 month	72	73
1970	< 1 month	34	44
	1 to 3 months	31	27
	>3 months	37	29

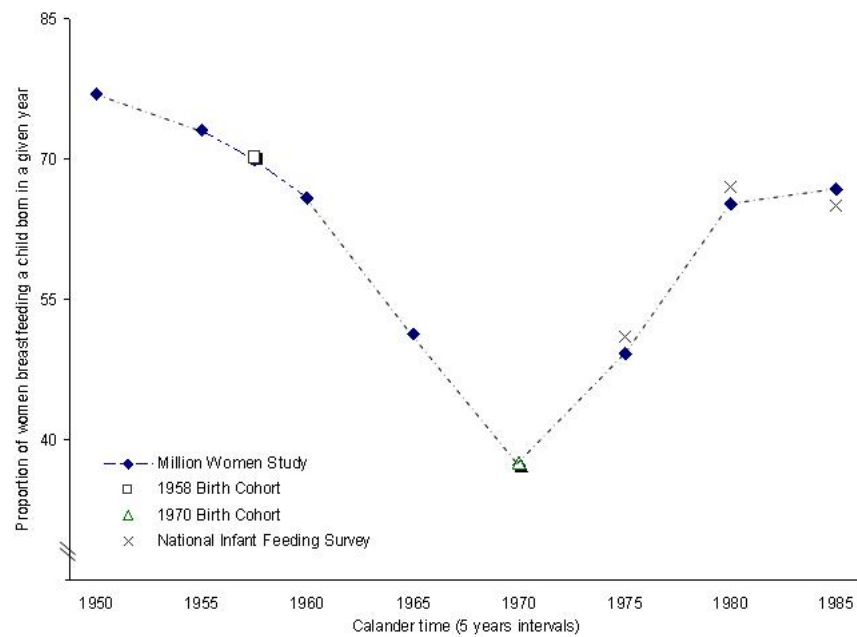


FIGURE 3.2: Trends in the proportion of UK women who breastfed over time using data from the Million Women Study, 1958 and 1970 birth cohorts, and the 1975, 1980, and 1985 National Infant Feeding surveys

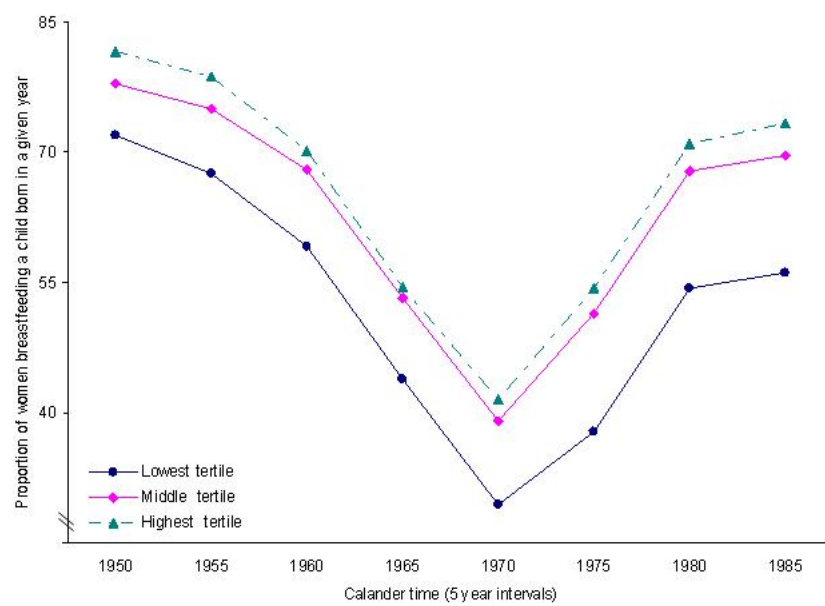


FIGURE 3.3: Trends in the proportion of UK women who breastfed over time using data from the Million Women Study stratified by socioeconomic tertile as measured at recruitment into the study

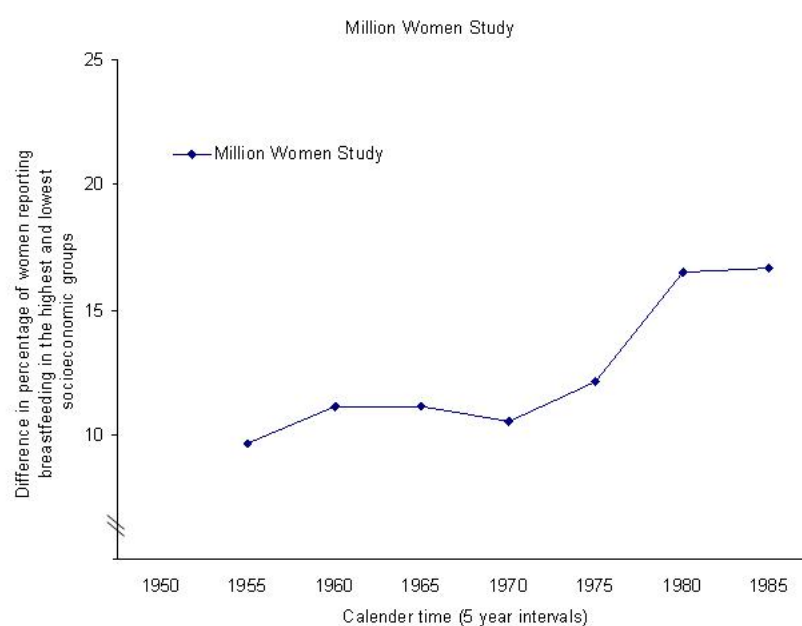


FIGURE 3.4: Data from the Million Women Study showing the difference in the proportion of women who breastfed over time between women in the highest socioeconomic groups and the lowest socioeconomic groups

Bibliography

- [1] Wright AL, Schanler RJ. The Resurgence of Breastfeeding at the End of the Second Millennium. *J Nutr.* 2001 Feb;131(2):421S–425.
- [2] Batty GD, Morton SMB, Campbell D, Clark H, Smith GD, Hall M, et al. The Aberdeen Children of the 1950s cohort study: background, methods and follow-up information on a new resource for the study of life course and intergenerational influences on health. *Paediatric and Perinatal Epidemiology.* 2004 May;18(3):221–239.
- [3] Institute of Education, University of London. for Longitudinal Studies C, editor. British Cohort Study [online]; 2010. Internet [cited 03 Dec 2010]. Available from: <http://www.cls.ioe.ac.uk/studies.asp?section=000100020002>.
- [4] Department of Health, editor. Breastfeeding & infant feeding [online]; 2010. Internet [cited 02 Dec 2010]. Available from: <http://www.dh.gov.uk/.../Breastfeedinginfantfeeding/index.htm>.
- [5] Power C, Elliott J. Cohort profile: 1958 British birth cohort (National Child Development Study). *Int J Epidemiol.* 2006 Feb;35(1):34–41.
- [6] Elliott J, Shepherd P. Cohort Profile: 1970 British Birth Cohort (BCS70). *Int J Epidemiol.* 2006 Aug;35(4):836–843.
- [7] Martin J, Monk J. Infant feeding 1975 : attitudes and practice in England and Wales : a survey carried out on behalf of the Department of Health and Social Security. London: H.M.S.O, Office of Population Censuses and Surveys, Social Survey Division; 1978.
- [8] Martin J, Monk J. Infant feeding 1980: a survey carried out on behalf of the Department of Health and Social Security and the Scottish Home and Health Department. London: H.M.S.O, Office of Population Censuses and Surveys, Social Survey Division; 1982.
- [9] Martin J, White A. Infant feeding 1985 : a survey carried out by Social Survey Division of OPCS on behalf of the Department of Health and Social Security and the Scottish Home and Health Department. London: H.M.S.O, Office of Population Censuses and Surveys, Social Survey Division; 1988.
- [10] Security D. Present day practice in infant feeding. In: Report on health and Social Subjects 9. London: HMSO; 1974. .
- [11] Martin J, Britain G, Britain G. Infant Feeding 1975: Attitudes and Practice in England And. London: HMSO; 1978.

- [12] Martin J, Monk J. Infant Feeding 1980: A Survey Carried Out on Behalf of The. London: OPCS; 1982.
- [13] Martin J, White A, Britain G. Infant Feeding 1985: A Survey Carried Out by Social Survey. London: HMSO; 1988.
- [14] Townsend P, Phillimore P, Beattie A. Health and deprivation. Routledge; 1988.
- [15] Institute of Education, University of London. Centre for Longitudinal Studies [online]; 2010. Internet [cited 03 Dec 2010]. Available from: <http://www.cls.ioe.ac.uk/default.asp>.
- [16] University of Essex. UK Data Archive [online]; 2011. Internet [cited 02 Feb 2011]. Available from: <http://www.data-archive.ac.uk/>.
- [17] Butland BK, Strachan DP, Lewis S, Bynner J, Butler N, Britton J. Investigation into the increase in hay fever and eczema at age 16 observed between the 1958 and 1970 British birth cohorts. *BMJ*. 1997;315(7110):717–721.
- [18] White A, Freeth S, O'Brien M. Infant feeding 1990: a survey carried out by the Social Survey Division of OPCS on behalf of the Department of Health, the Scottish Home and Health Department, the Welsh Office and the Department of Health and Social Services in Northern Ireland. H.M.S.O, Office of Population Censuses and Surveys Social Survey Division; 1992.
- [19] Foster K, Lader D, Cheesbrough S. Infant feeding 1995 : a survey of infant feeding practices in the United Kingdom carried out by Social Survey Division of ONS on behalf of the Department of Health, the Scottish Office Department of Health, the Welsh Office and the Department of Health and Social Services in Northern Ireland. London: Stationery Office, Government Statistical Service; 1997.
- [20] Apple RD. The medicalization of infant feeding in the United States and New Zealand: two countries, one experience. *Journal of Human Lactation: Official Journal of International Lactation Consultant Association*. 1994 Mar;10(1):31–37.
- [21] Fomon S. Infant feeding in the 20th century: formula and beikost. *The Journal of Nutrition*. 2001 Feb;131(2):409S–205.
- [22] Ryan KM, Grace VM. Medicalization and women's knowledge: the construction of understandings of infant feeding experiences in post-WW II New Zealand. *Health Care for Women International*. 2001 Aug;22(5):483–500.
- [23] Foss KA. Perpetuating "scientific motherhood": infant feeding discourse in *Parents magazine*, 1930-2007. *Women & Health*. 2010 May;50(3):297–311.

Chapter 4

Persistent effects of childbearing on body mass index in postmenopausal UK women: cross-sectional analysis in a large population

1

¹Work from this chapter has been presented and published in conference proceedings. [1] In addition, the findings have been written up as a scientific paper and are in the process of being submitted for publication.

Abstract

Objectives: *To explore the association between childbearing patterns and women's body mass index (BMI) after the menopause*

Design: *Cross-sectional analysis of the relationship between reproductive history and BMI in postmenopausal women*

Setting: *Population-based study of UK women recruited in 1996 to 2001 through the National Health Service Breast Screening Programme (the Million Women Study)*

Participants: *740 628 women who at recruitment were postmenopausal and who reported their reproductive histories, anthropometry, and other personal and socio-demographic factors*

Main outcome measure: *Mean BMI (kg/m^2) at recruitment in relation to parity and breastfeeding history, standardised by age, region of residence, socioeconomic status, smoking, and physical activity*

Results: *Women were aged 57.5 years [SD 4] on average, and had a standardised mean BMI of $26.2 \text{ kg}/\text{m}^2$ [SD 5]; 88% were parous, with mean parity of 2.1 children [SD 1.2]. Among parous women 70% had ever breastfed and the average total duration of breastfeeding was 4.7 months [SD 7.8]. The standardised mean BMI increased progressively with number of births from $25.6 \text{ kg}/\text{m}^2$ in nulliparous women up to $27.2 \text{ kg}/\text{m}^2$ for women with four or more births, a difference of 1.7 (95% confidence interval 1.6 to 1.7) kg/m^2 . At every parity level the standardised mean BMI was lower among women who had breastfed. The mean BMI decreased the longer women had breastfed and was 0.5 (0.5 to 0.6) kg/m^2 lower in those who had breastfed for a total of 10 months or more (average 18 months) than in women who had never breastfed. These differences were highly statistically significant ($p < 0.0001$) and independent of the effects of socioeconomic group, region, smoking, and physical activity. The associations between parity, breastfeeding and BMI in postmenopausal women were of a similar order of magnitude to established risk factors known to be associated with BMI, for example smoking, and physical activity.*

Conclusions: *Among postmenopausal women in the UK the mean BMI increases with the number of children women have had, but decreases the longer they had breastfed their children.*

4.1 Introduction

Excess body weight is an established and important risk factor for vascular disease and, to a lesser extent, for cancer. [2, 3] For example, in a recent analysis of prospectively collected individual data in almost one million adults, all cause mortality and mortality from a range of specific causes increased with increasing body mass index (BMI) over and above 25 kg/m². [2]

Among reproductive-aged women in developed countries weight, in the short-term, tends to increase with each birth, [4–6] but the published evidence of the short-term effects of breastfeeding on body weight is inconsistent. [7, 8] In the longer term, after a woman's reproductive years are over, some but not all investigators have reported that women's adiposity (variously measured) increases with the number of children they have had (most have reported on the association between parity and the risk of women being overweight or obese, defined by WHO as a BMI above 25 kg/m² and above 30 kg/m², respectively). [9–14] [15] To my knowledge only two small studies have looked at the effect of breastfeeding history on postmenopausal adiposity and their findings are unclear. [16, 17]

Women's level of adiposity as captured by BMI is known to be related to socioeconomic factors, smoking, and physical activity, and these factors are associated with reproductive history. [18–24] In this chapter I report on the association between women's childbearing histories and BMI after the menopause taking into account the effects of important co-factors.

4.2 Methods

4.2.1 Study Population

Recruitment

The Million Women Study is a prospective study of women aged between 50 and 64 years when they were invited for screening by the National Health Service (NHS) Breast Screening Programme in England and Scotland between 1996 and 2001. A total of 1.3 million women completed a study questionnaire at recruitment which asked about their height, weight, reproductive history, socioeconomic and lifestyle factors, and other personal characteristics.

Details of the study design and methods relevant to this thesis are described more fully in chapter 2 (page 18), and have been published previously. [25] Study questionnaires can be viewed at <http://www.millionwomen.org> and have been included in this thesis in Appendix A, page 241. The study was approved by the Anglia and Oxford Multi-Centre Research Ethics Committee. All participants gave their written consent to take part in the study.

4.2.2 Data Definitions

Menopause

Women were classified as postmenopausal if they reported their periods stopped naturally or after bilateral oophorectomy, or if they were aged 53 years or older (most women aged 53 years who had a natural menopause were postmenopausal by that age).

Parity and breastfeeding

At recruitment women were asked to report the number of children they had given birth to (including still births), which was used to define parity. Women were asked to report, for each birth, if they had breastfed and if so, the duration of breastfeeding in months. This information was used to define breastfeeding (ever or never), and total duration of breastfeeding (summation over all children of reported duration of breastfeeding in months). In addition to using women's total duration of breastfeeding I also calculated the average duration of breastfeeding per child using each woman's total duration of breastfeeding divided by her parity (summarised into per child breastfeeding duration <6 months, or ≥ 6 months). No questions were asked about the exclusivity of breastfeeding.

Body mass index (BMI)

Women's reported current weight and height at recruitment were used to derive body mass index (BMI) as $\text{weight (kg)} / \text{height (m)}^2$. In addition to BMI calculated from self-reported data I used GP-measured heights and weight data that were available for a random sample of 3,745 participants. In these analyses BMI was modeled as a continuous and a categorical variable. Where appropriate I divided BMI into the following categories <22.5 , 22.5 to 24.9, 25 to 27.4, 27.5 to 29.9, 30 to 32.4, 32.5 to 35, or $\geq 35 \text{ kg/m}^2$, and 22.5 to 24.9 kg/m^2 was used as the reference category.

Co-factors & effect modifiers

Participants in the Million Women Study are classified into 10 geographic regions of residence within England and Scotland. Women's socioeconomic status was summarised into fifths (based on the Townsend deprivation index). [26] Women's smoking status was categorised as never, past or current smokers. Current smokers were further sub-divided by the number of cigarettes smoked per day (current <15 cigarettes per day, or current ≥ 15 cigarettes per day). Women reported the number of times per week they engaged in any physical activity, I summarised this into three categories (≤ 1 time per week, 2 to 3 times per week, ≥ 4 times per week).

4.2.3 Statistical Analysis

Eligibility and exclusion criteria

Women were eligible for inclusion into the main analysis if they were postmenopausal, had not been registered with a cancer (other than non-melanoma skin cancer) before recruitment, and had been asked about their breastfeeding histories. Of the 936,102 women who were initially eligible I excluded women with missing or incomplete information on parity (0.2%), breastfeeding (10%), BMI (4%), socioeconomic group (0.7%), physical activity (1.4%), and smoking (4.4%). Thus the analyses were restricted to postmenopausal women with complete information on the above variables. The final number of women included in the main analysis was 740,628 (about 80% of those eligible). Figure 4.1 (page 107) shows a flow diagram of the eligibility and inclusion criteria and the numbers of women excluded.

Exposures, outcome, and covariates

The main exposure variables were parity and duration of breastfeeding modeled as categorical variables. Parity was summarised as 0, 1, 2, 3, or 4 or more; and total duration of breastfeeding as never breastfed, breastfed for less than 6 months, breastfed for 6 to 9 months, or breastfed for 10 months or more (these approximate quartiles of total duration for women who breastfed). The main outcome in this analysis was BMI calculated using women's self-reported height and weight data, which was modeled as a continuous variable. The following co-factors were included in statistical models for the main analyses: age (continuous variable), region of residence (10 regions in the UK), quintiles of socioeconomic status, categories of smoking, and physical activity.

Model specifics

I used Spearman's correlation and Bland-Altman graphical analysis techniques to assess the agreement between BMI calculated from self-reported data and measured data. [27, 28] For the main analysis I used linear regression to estimate the mean change in BMI associated with increasing parity and duration of breastfeeding, with adjustment for all of the factors listed above. The mean BMI, standardised by the variables listed above, was estimated using the regression coefficients by category of parity, breastfeeding, socioeconomic group, smoking and physical activity. I fitted parity and duration of breastfeeding as continuous variables in tests for trend (using the category means). For results presented as figures standardised mean BMI (with 95% confidence intervals) are shown for categories of parity, breastfeeding and, where appropriate, other factors.

Additional analyses

It is known that BMI calculated from self-reported data may be subject to misclassification errors. [3] To assess the potential effects of measurement error on model estimates due to the use of BMI calculated using self-reported height and weight data I repeated the main analysis restricting the study population to women who had BMIs calculated using GP-measured height and weight data (of the 3,745 randomly selected women with BMIs calculated from both self-reported and measured data 2,800 were eligible for inclusion in this analysis). In addition I used the regression dilution technique; [29–31] I categorised BMIs from self-reported height and weight data as previously described but I calculated the average for each category using data from the subgroup of women with measured height and weight data. I applied these categories to the study population used for the main analysis. I then repeated the main analyses using the category means and modeling BMI as a continuous outcome. All analyses were performed using STATA version 10 (Stata corporation, TX, USA).

4.3 Results

Characteristics of the study population

In total, 740,628 postmenopausal women were included in the analysis (80% of those eligible). The mean age of the women was 57.5 years (SD 4) and 88% reported having had at least one child, with mean parity 2.1 (SD 1.2). The mean age at first birth was 23.9 years (SD 4) and the mean age at last birth was 28.5 years (SD 5). Among parous women, 70% had breastfed. The mean total duration of breastfeeding was 4.7 months (SD 7.8) and increased with increasing parity. The mean duration of breastfeeding per child was 1.9 months (SD 2.9) and increased slightly with increasing parity.

Table 4.1 (page 108) shows characteristics of study participants by parity and, among parous women, by breastfeeding history. Nulliparous women tended to be of a higher socioeconomic status, and were less likely to be current smokers, and more likely to report engaging in regular physical activity than parous women. Among parous women, increasing parity was associated with lower socioeconomic status, an increasing likelihood of being a current smoker, and a lower likelihood of engaging in regular physical activity. Women who had breastfed were of a higher socioeconomic status, less likely to be current smokers, and more likely to engage in regular physical activity than women who had not breastfed.

Figure 4.2 (page 109) shows the agreement between BMIs calculated from self-reported height and weight data, and BMIs calculated from measured height and weight data. There was strong association between BMIs calculated from self-reported data and BMIs calculated from measured data (Spearman's correlation coefficient = 0.95). Figure 4.3 (page 109) shows the Bland-Altman plot with the average of BMI calculated from

measured and self-reported data on the x-axis and the mean difference between these measures plotted on the y-axis. There is some evidence to suggest that BMIs calculated from women's self-reported data underestimate women's true BMI as calculated from measured heights and weights and this was more likely among women with a BMI greater than 30. However the mean difference between measured and self-reported BMIs was small and not significantly different from zero (-0.65 kg/m^2 , CI -3.10 to 1.79 kg/m^2).

Univariate analyses

Table 4.2 (page 110) shows the effect on women's BMI after the menopause by parity, breastfeeding, socioeconomic group, smoking, and physical activity in models adjusted for age and region only. In these minimally adjusted analyses parous women had a higher BMI than nulliparous women. Among parous women, those who breastfed had a lower BMI than those who never breastfed. BMI differed according to socioeconomic group; women in the highest socioeconomic tertile had a BMI that was more than one BMI unit lower than women in the lowest tertile. Women who had never smoked had a higher BMI than women who reported being current smokers. Increasing physical activity was also associated with a lower BMI.

Multivariate analyses

Table 4.3 (page 111) shows the change in mean BMI with increasing parity (adjusted for total duration of breastfeeding) and with increasing total duration of breastfeeding (adjusted for parity) standardised by age and region only, or for the co-factors included in the main analyses (socioeconomic group, smoking, and physical activity). Although

standardisation by these factors slightly attenuated the results, there was still an independent effect of increasing parity and increasing duration of breastfeeding on standardised mean BMI.

Table 4.3 (page 111) shows the change in mean BMI associated with childbearing in three models. In the left column the linear regression model of the association between parity or breastfeeding and BMI is standardised only for age and region (Model A). In the middle column the regression model is additionally standardised by breastfeeding (when considering parity), and for parity (when considering breastfeeding). In the right-hand column the final model is presented, which is Model A with additional standardisation for parity, breastfeeding, socioeconomic group, smoking, and physical activity.

It can be seen from the table that in Model A increasing parity was associated with an increase in mean BMI, while an increasing duration of breastfeeding was associated with a modest decrease in mean BMI. Standardising Model A by breastfeeding (when considering parity), and by parity (when considering breastfeeding) strengthened the positive association between parity and mean BMI and the negative association between duration of breastfeeding and BMI. It can be seen that although the associations between parity and breastfeeding and mean BMI were slightly attenuated in the final model (which was standardised for age, region, parity, breastfeeding, socioeconomic group, smoking, and physical activity) there was still an independent effect of increasing parity and increasing duration of breastfeeding on standardised mean BMI. Table 4.3 (page 111) also shows the associated change in mean BMI per 6 months of breastfeeding using category means for the trend analysis. The association between parity and mean BMI was non-linear it was therefore not appropriate to fit a linear trend for parity in this analysis.

Figure 4.4 (page 112) is similar to table 4.3 (page 111) and shows the mean BMI in middle-age by parity and breastfeeding history, standardised by age, region, socioeconomic group, smoking, and physical activity. Again, women who never had children had a lower mean BMI than parous women. It can be seen in figure 4.4 (page 112) that both parity and breastfeeding history have an effect on BMI in postmenopausal women. The standardised mean BMI among parous women increased with each additional child; with a relatively small increase between women with one or with two children, and a more marked increase for parity greater than two. For any given parity, women who had ever breastfed had a significantly lower standardised mean BMI than women who had never breastfed.

Figure 4.5 (page 113) shows the standardised mean BMI according to parity and to breastfeeding status separately for women in upper, middle, and lower socioeconomic tertiles. At every level of socioeconomic status, there was an independent effect of parity and breastfeeding on BMI. The standardised mean BMI was consistently lower in higher socioeconomic groups, for example in nulliparous women the standardised mean BMI was 25.3 kg/m² in the highest socioeconomic group and was significantly higher ($p < 0.0001$), at 26.4 kg/m², in the lowest group (a difference of 1.1 (CI 1.0 to 1.2) kg/m²). There was a positive association between parity and standardised mean BMI and irrespective of socioeconomic group and parity, women who breastfed had a lower standardised mean BMI than women who did not. For example, women of parity three in the highest, middle, or lowest socioeconomic tertile who breastfed had a standardised mean BMI of 25.8 kg/m², 26.1 kg/m², and 26.7 kg/m² respectively; in comparison women who did not breastfeed had a standardised mean BMI of 26.1 kg/m², 26.7 kg/m², and 27.2 kg/m² respectively.

Figure 4.6 (page 114) illustrates the effect of increasing duration of breastfeeding on the

relationship between parity and standardised mean BMI. Again it can be seen that parity and breastfeeding independently affect BMI. It can also be seen that at every level of parity, the standardised mean BMI is lower the longer the duration of breastfeeding per child. For example, among women of parity three, those who did not breastfeed had a standardised mean BMI of 26.7 kg/m², those who breastfed each child for up to 6 months had a standardised mean BMI of 26.2 kg/m² (CI), while those who breastfed each child for 6 months or more had a standardised mean BMI of 25.9 kg/m². Using the average duration of breastfeeding per child is equivalent to results obtained from a model that includes parity and total duration of breastfeeding. The purpose of representing the data in this way is to demonstrate on a more appropriate scale the lengths of total duration of breastfeeding required at various parity levels to achieve the change in mean standardised BMI the model predicts. For example, women of parity three would have a mean standardised BMI equivalent to women of parity one who breastfed for less than 6 months if they breastfed each child on average 6 months or more, in other words if their total duration of breastfeeding was 18 months or more.

Findings in context

Figure 4.7 (page 115) shows the standardised mean BMI by socioeconomic group, smoking, and physical activity as well as by parity and total duration of breastfeeding (for each characteristic the mean BMI is standardised for all other factors shown in the figure). It can be seen that each characteristic has an independent effect on BMI in this population of postmenopausal women. Comparing the relative magnitude of the effects of each factor on standardised mean BMI, women with four or more children have a mean BMI 1.7 kg/m² greater (CI 1.6 to 1.7) than nulliparous women, and parous women with a total duration of breastfeeding of 10 months or more (on average 18

months) have a mean BMI 0.5 kg/m^2 lower (CI 0.5 to 0.6) than women with similar characteristics who had not breastfed their children. The difference in the standardised mean BMI between women in the lowest and highest socioeconomic quintiles was 1.2 kg/m^2 (CI 1.1 to 1.2), and between never and current smokers was 1.1 kg/m^2 (CI 1.0 to 1.1). Women who reported engaging in physical activity four or more times per week had a standardised mean BMI that was 1.8 kg/m^2 lower (CI 1.7 to 1.8) than women who engaged in physical activity less than once per week.

Additional analyses

Effect on model estimates of restricting the study population to women with BMI calculated using measured height and weight data

Table 4.4 (page 116) presents the previous models fitted using only the 2,800 women with measured heights and weights who were eligible for inclusion into this analysis (0.4% of the study population). As compared to the results shown in table 4.3 (page 111) it can be seen that the point estimates have changed for some categories; the estimates for the effects of parity on standardised mean BMI are reduced whereas the estimates for the effects of breastfeeding are slightly larger, the most marked change is in the widening of the confidence intervals.

Effect on model estimates of using measured heights and weights to calculate BMI and the regression dilution technique to correct for misclassification

I repeated this analysis using the mean measured BMI data as the average for categories of self-reported BMI and applied to the study population used for the main analysis. Table 4.5 (page 117) shows the results of this analysis. Comparing these findings with

results from the main analysis presented in table 4.3 (page 111) it can be seen that the main findings are unchanged - the magnitude of the effect of childbearing on BMI after the menopause was similar and statistically significant. Increasing parity was associated with an increase in the standardised mean measured BMI, and standardised mean measured BMI was lower with longer durations of breastfeeding. In addition there was little to suggest that using self-reported BMI biased the results to any great degree.

4.4 Discussion

Summary

In this study of almost 750,000 postmenopausal women in the UK aged 57.5 years on average, there were persistent and independent effects of parity, and of breastfeeding history on BMI in later life. These associations were also independent of other known important risk factors - socioeconomic group, smoking, and physical activity.

Findings in relation to other studies

The association between increasing parity and increased risk of obesity in middle-aged women has been reported previously but this is the largest study conducted so far. [10–14] These findings extend the evidence and show that the effect is similar in magnitude to some of the known important risk factors for increased BMI. In addition, I found that breastfeeding has a long-term beneficial effect on BMI in postmenopausal women, offsetting to some extent the increase in BMI associated with increasing parity. To date only two small studies have reported on the relationship between breastfeeding history and maternal BMI or a related measure in later life. Among middle-aged

Swedish women increasing duration of lactation was associated with decreased risk of abdominal obesity; [16] and in another study, also of middle-aged European women (born between August 1944 and April 1946 during the Dutch famine and who may be a somewhat atypical population), BMI was lower among women who had breastfed than among those who had not. [17] Both these findings are consistent with our results. The relationship between childbearing and BMI in postmenopausal women is in part due to the effect of parity and in part due to the effect of breastfeeding. Since breastfeeding reduces BMI, studies that do not also account for women's breastfeeding histories will tend to slightly underestimate the effect of each birth on women's BMI.

Possible explanations

The relationships between parity, breastfeeding and BMI found here are unlikely to be due to chance given the very small p-values. They are also unlikely to be the result of reverse causation as childbearing and breastfeeding occurred on average 30 years before women's weights and heights were recorded in this study. In this study socioeconomic status, smoking, and physical activity were all related to BMI. However, when I adjusted for these factors the association between parity, breastfeeding and BMI was only slightly attenuated, and in mutually adjusted and stratified analyses the associations between BMI and parity and breastfeeding were shown to be largely independent of those between BMI and the other risk factors. The associations between BMI and socioeconomic group, smoking, and physical activity in this study are of a broadly similar magnitude to those reported in the published literature. [19–23] [24]

Strengths and weaknesses

These findings should be interpreted within the context of the study design. Parity and breastfeeding history were obtained by self-report, but studies in older women have found that their long-term recall of these events is reliable. [32–35] BMI was calculated using women's self-reported heights and weights and may be affected by random and systematic measurement error, however previous studies within similar population groups have shown that this is unlikely to be a source of major bias. [36, 37] In addition I found a high degree of correlation when I compared BMI calculated from measured weight and height with self-reported data. Women excluded from the main analyses because of missing data were of a similar age, parity, breastfeeding status, and socioeconomic status to those included, however they were more likely to be current smokers and less likely to engage in regular physical activity.

Strengths of this study include its size and the availability of information about co-factors. The Million Women Study includes one in four UK women aged between 50 and 64 years at the time of recruitment. Detailed information on socio-economic factors, smoking and physical activity allowed for fine subdividing by key co-factors whilst retaining sufficient power to assess the effects of childbearing and breastfeeding on BMI. The study population is ethnically homogeneous with 98% reporting they were of white ethnicity.

Conclusions

These findings and those from the published literature relate to postmenopausal women in developed countries, and do not necessarily apply to women in other settings, particularly in developing countries. For women in the UK, and probably also women

in other developed countries, these results provide good evidence that even after the menopause, reproductive history has a long-lasting effect on BMI. I also found that the magnitude of effect of these factors on BMI is broadly similar in range to established risk factors such as socioeconomic group, smoking, and physical activity. For example, the difference in standardised mean BMI between women with 4 or more children versus women with a parity of one is similar in size to the difference in standardised mean BMI between women in the highest versus the lowest socioeconomic quintiles. These and other published findings suggest that, when younger women are counseled on infant feeding choices, attention should be given to the benefits breastfeeding may have on their own health (which may include benefits in terms of long-term risk of excess adiposity), as well as to the benefits on the health of their infants.

4.5 Chapter summary

In this chapter I explored the association between parity and breastfeeding and women's BMI after the menopause. I found persistent, independent, and dose-related effects for both parity and breastfeeding history on maternal BMI after the menopause. As BMI is known to be somewhat associated with socioeconomic group, and socioeconomic group is also known to be somewhat associated with childbearing I looked carefully at how socioeconomic group affected the association found between childbearing and postmenopausal BMI. I found good evidence that the association between childbearing and maternal BMI after the menopause is independent of socioeconomic group.

Research into possible biological mechanistic processes which may underpin these associations is limited. Among women in their reproductive years small prospective studies of post-partum women have shown variable short-term effects of childbearing on post-partum weight trajectories, [38, 39] maternal glucose metabolism, lipid profile, and blood pressure. [40–45] For parity, possible explanations include pregnancy as a diabetogenic state, [46] hormonal effects of fewer ovulatory cycles, [47] and the lifestyle associated with raising children. [48] For breastfeeding, possible explanations include increased metabolic demand associated with breastfeeding, and the effects of lactation associated hormones (prolactin and oxytocin) in “resetting maternal metabolism”. [49]

As discussed above these findings have public health implications, they also have implications for subsequent analyses explored in this thesis; it is known that raised BMI is an important risk factor for diabetes mellitus and for heart disease. In exploring whether childbearing is associated with women's risk of diabetes mellitus, or ischaemic heart disease after the menopause it will be important to consider whether this is independent of the effect of childbearing on BMI or whether the associations are largely

mediated through an effect of childbearing on women's BMI.

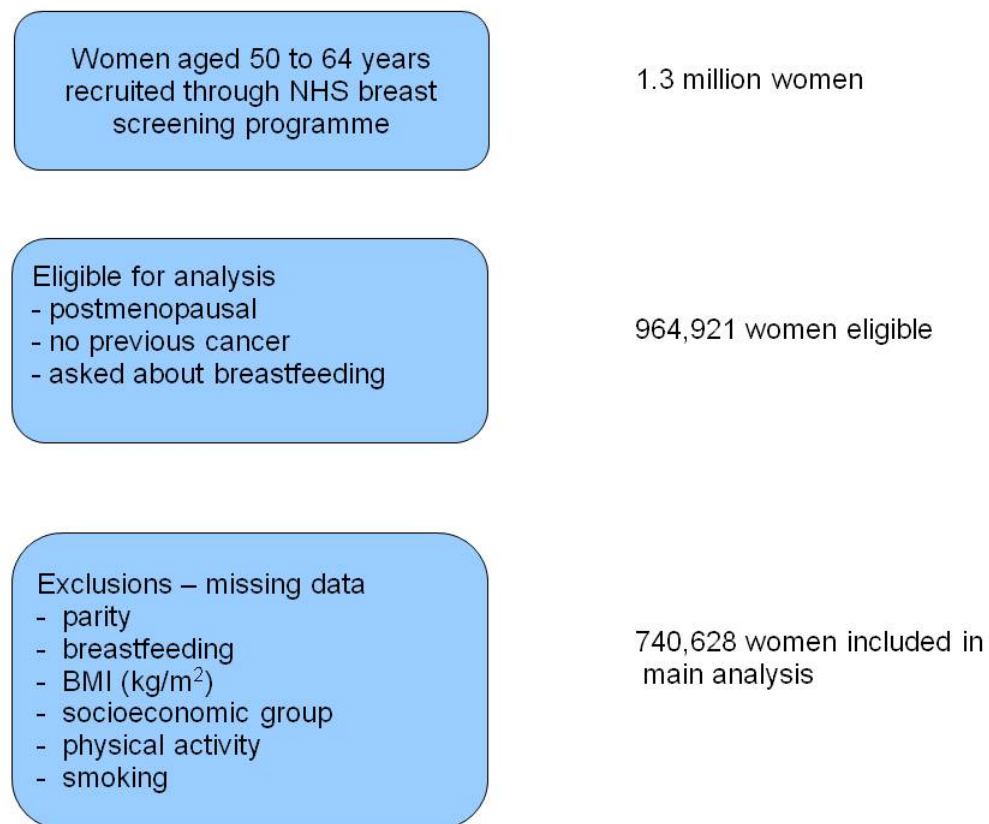


FIGURE 4.1: Flow diagram showing eligibility and inclusion criteria and the final number of women included in the main analyses

TABLE 4.1: Characteristics of study participants used in the analysis of body mass index by parity and breastfeeding status

Population Characteristics	By Parity					By Breastfeeding status*		All Women
	Nulliparous n = 89 448	One n = 100 639	Two n = 310 841	Three n = 159 100	Four or more n = 80 600	Never Breastfed n = 201 688	Ever Breastfed n = 449 492	n = 740 628
Age in years (Mean [SD])	57 (5)	57 (5)	57 (4)	58 (4)	58 (4)	57 (4)	58 (4)	57 (4)
Parity (Mean [SD])	0	1 (0)	2 (0)	3 (0)	4.4 (0.8)	2.1 (0.9)	2.5 (1.0)	2.1 (1.2)
Number of children breastfed (Mean [SD])	0	0.5 (0.5)	1.1 (0.9)	1.7 (1.2)	2.5 (1.8)	0	1.9 (1.0)	1.2 (1.2)
Women who ever breastfed (%)	0	52	68	76	80	0	100	70
Duration breastfeeding								
Total duration in months (Mean [SD])	0	1.9 (3.5)	4.3 (6.2)	6.8 (8.7)	10.5 (13.9)	0	7.7 (8.8)	4.7 (7.8)
Duration per child in months (Mean [SD])	0	1.9 (3.5)	2.1 (2.9)	2.3 (2.9)	2.3 (3.0)	0	3.1 (3.1)	1.9 (2.9)
In lower third of socioeconomic group (%)	33	34	30	36	50	38	32	34
Current smokers (%)	17	20	17	19	25	23	17	19
Physical activity less than once a week (%)	26	30	27	28	34	34	26	28

* Among parous women

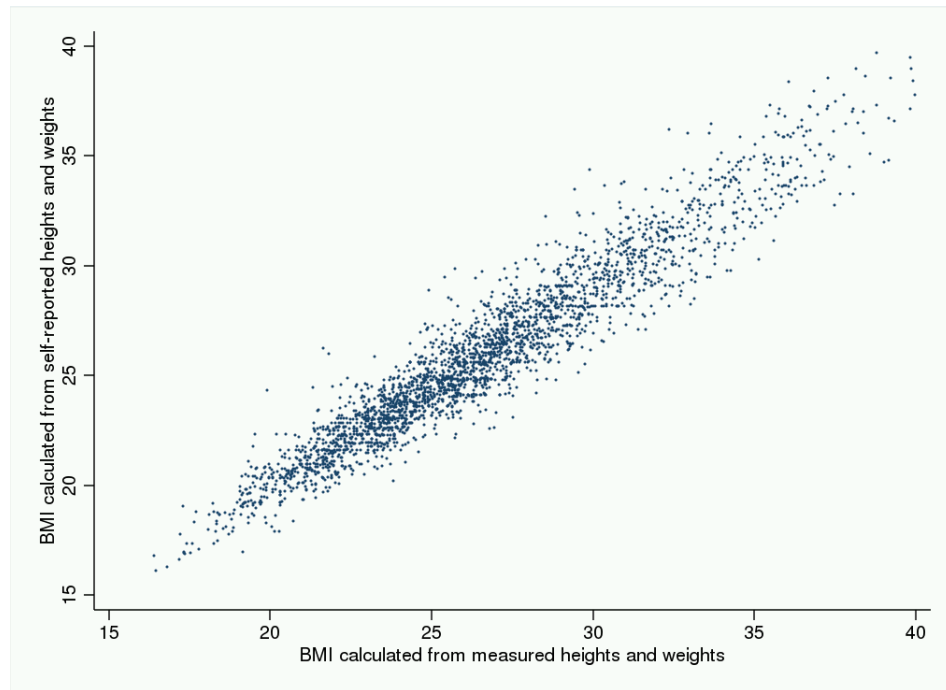


FIGURE 4.2: Scatter plot showing the association between BMIs calculated from measured height and weight data taken from a random sample of Million Women Study participants (x-axis) and self-reported height and weight data from the recruitment questionnaire (y-axis)

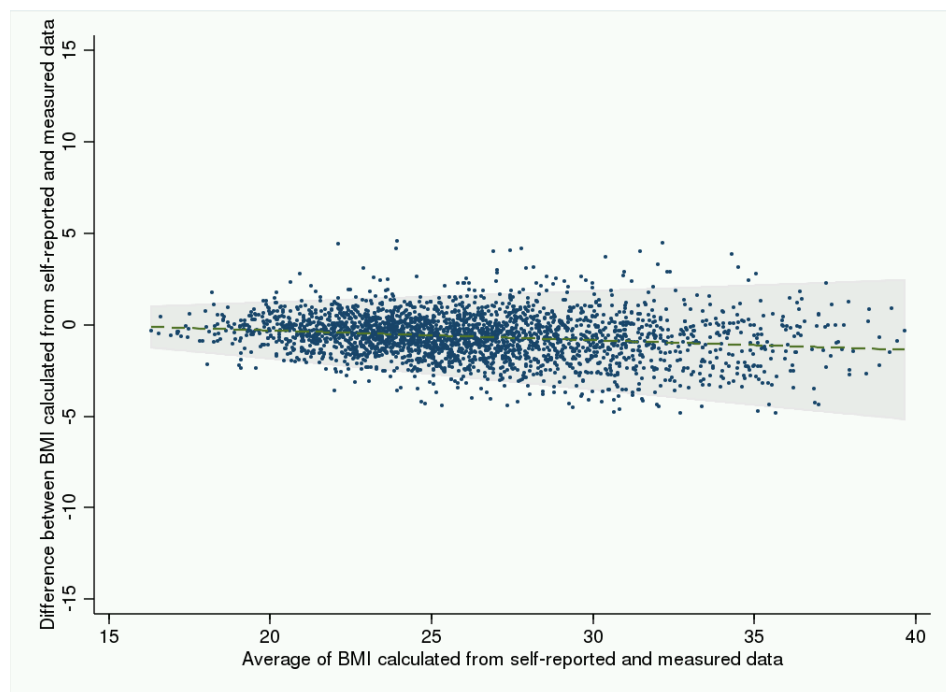


FIGURE 4.3: Bland-Altman plot showing relationship between the average BMI calculated from the two data sources (x-axis) and the difference between the BMIs calculated from the two data sources (y-axis)

TABLE 4.2: Association between parity and breastfeeding and mean BMI (kg/m^2) among postmenopausal women in a model adjusted only for age and region

Characteristics of women		Number of women	Mean BMI (kg/m^2)	(95% CI)
Parity				
	Parous	651,180	26.28	(26.27 - 26.29)
	Nulliparous	89,448	25.56	(25.53 - 25.60)
LR test heterogeneity				$p < 0.0001$
Breastfed				
	No	291,136	26.40	(26.38 - 26.41)
	Yes	449,492	26.06	(26.05 - 26.08)
LR test heterogeneity				$p < 0.0001$
Socioeconomic group				
	Upper third	244,575	25.68	(25.67 - 25.70)
	Middle third	243,550	26.09	(26.07 - 26.11)
	Lowest third	252,503	26.79	(26.77 - 26.81)
LR test heterogeneity				$p < 0.0001$
Smoking				
	Past	214,953	26.70	(26.68 - 26.72)
	Never	386,641	26.12	(26.14 - 26.17)
	Current <15	72,345	25.21	(25.18 - 25.25)
	Current >15	66,689	25.82	(25.78 - 25.85)
LR test heterogeneity				$p < 0.0001$
Any physical activity				
	< 1 per week	321,103	27.06	(27.05 - 27.08)
	≥ 1 per week	419,525	25.53	(25.52 - 25.54)
LR test heterogeneity				$p < 0.0001$

TABLE 4.3: Change in mean BMI (kg/m^2) among postmenopausal women among parous women by parity and total duration of breastfeeding in models variously standardised

Unadjusted Results		Model A - standardised by age and region only	Model A + additionally standardised by breastfeeding (when looking at parity), and for parity (when looking at breastfeeding)	Model A + additionally standardised by breastfeeding, parity, socioeconomic group, smoking, and physical activity
Parity (mean)	Number of women			
1	100,639	reference	reference	reference
2	310,841	-0.02 (-0.05 to 0.01)	-0.01 (-0.04 to 0.02)	0.13 (0.09 to 0.16)
3	159,100	0.45 (0.41 to 0.48)	0.45 (0.41 to 0.48)	0.58 (0.55 to 0.62)
≥ 4 (4.4)	80,600	1.31 (1.27 to 1.36)	1.30 (1.26 to 1.34)	1.33 (1.28 to 1.37)
Total duration of breastfeeding (mean in months)				
Did not breastfeed	201,688	reference	reference	reference
<6 months (2.3)	239,836	-0.27 (-0.24 to -0.30)	-0.26 (-0.24 to -0.29)	-0.24 (-0.21 to -0.26)
6 - 9 months (7.3)	82,198	-0.43 (-0.40 to -0.47)	-0.42 (-0.38 to -0.46)	-0.36 (-0.32 to -0.40)
≥ 10 months (18.5)	127,458	-0.44 (-0.41 to -0.48)	-0.43 (-0.40 to -0.46)	-0.53 (-0.50 to -0.57)
Change in mean BMI per 6 months breastfeeding*		-0.12 (-0.11 to -0.13)	-0.11 (-0.10 to 0.12)	-0.13 (-0.11 to -0.13)

* trend fitted using continuous duration of breastfeeding data (per month) and then multiplied by six to get the trend per six months

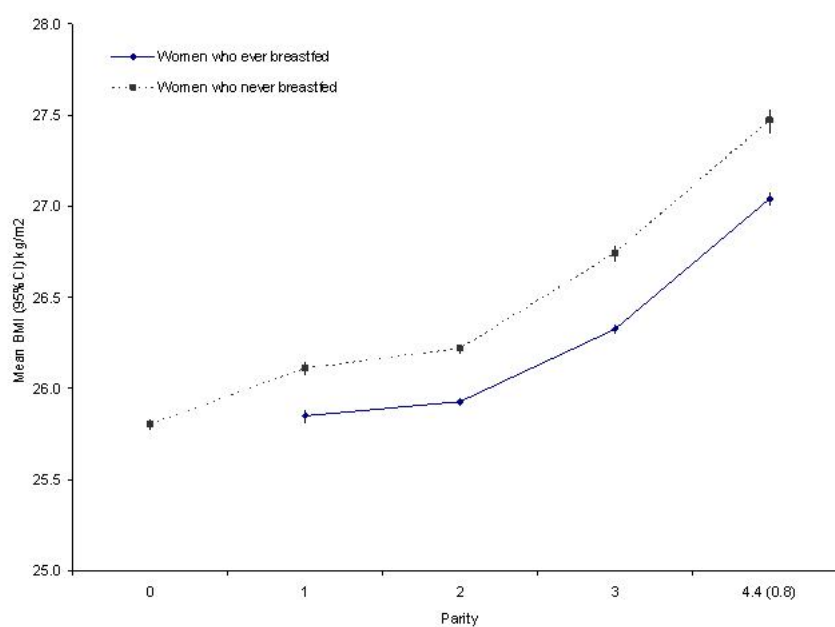


FIGURE 4.4: Mean BMI (kg/m^2) among postmenopausal women by parity and breastfeeding standardised by age, region, socioeconomic group, smoking, and physical activity

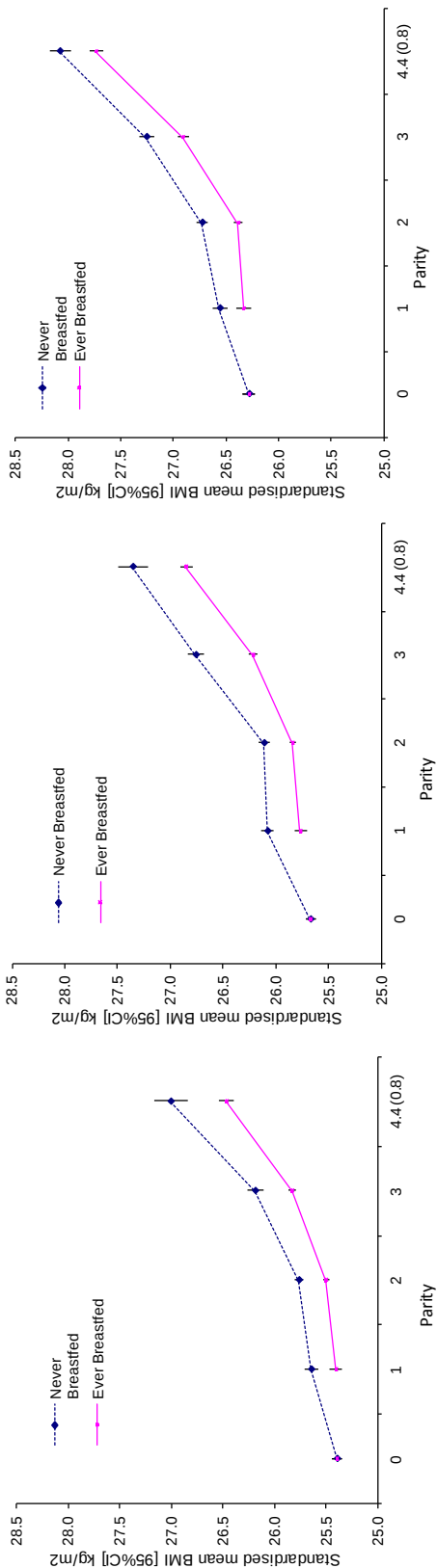


FIGURE 4.5: Mean BMI (kg/m^2) among postmenopausal women by parity and breastfeeding standardised by age, region, socioeconomic group, smoking, and physical activity in upper, middle and lower socioeconomic tertiles

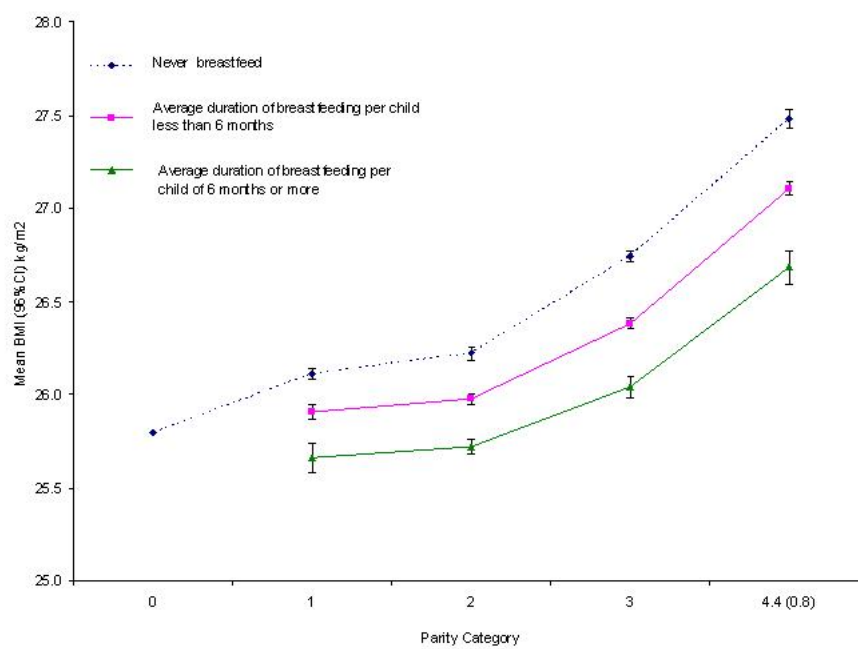
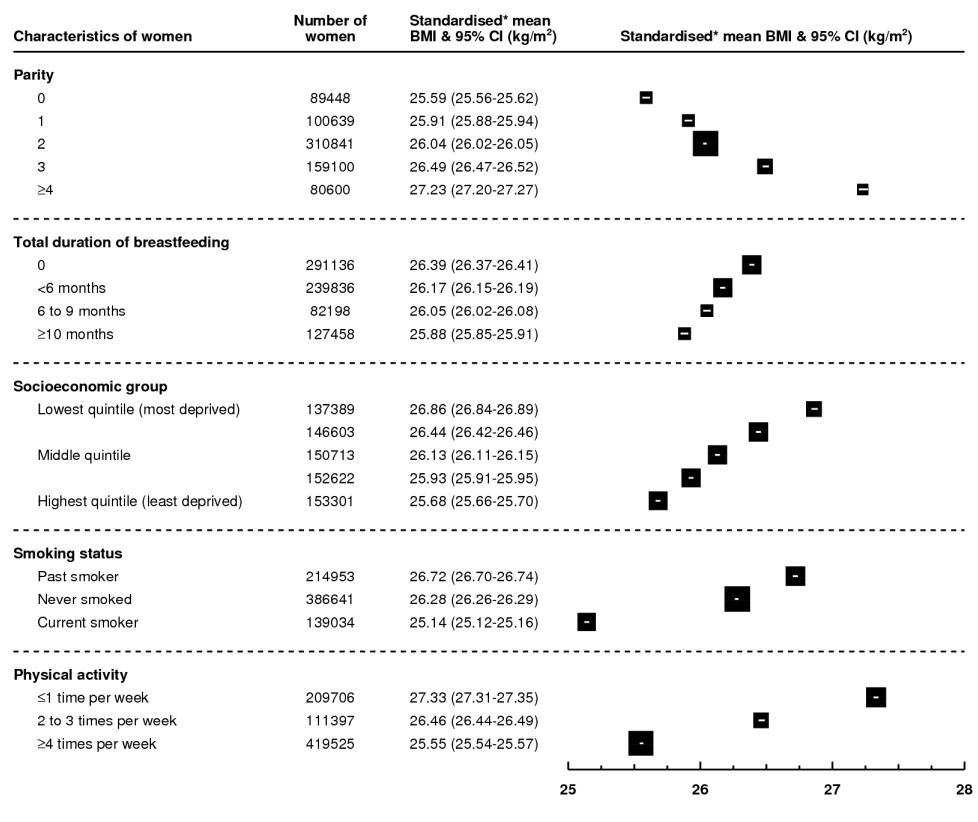


FIGURE 4.6: Mean BMI (kg/m^2) among postmenopausal women by parity and increasing average duration of breastfeeding per child standardised by age, region, socioeconomic group, smoking, and physical activity



* standardised by age, region, parity, total duration of breastfeeding, socioeconomic group, smoking status and physical activity, as appropriate.

FIGURE 4.7: Change in standardised mean BMI (kg/m²) among postmenopausal women for selected characteristics

TABLE 4.4: Change in standardised mean BMI (kg/m^2) among postmenopausal women by parity and total duration of breastfeeding for a randomly selected sample of 2,800 women with BMIs calculated from measured heights and weights

		Model A - standardised by age and region only	Model A + additionally standardised by breastfeeding (when looking at parity), and for parity (when looking at breastfeeding)	Model A + additionally standardised by breastfeeding, parity, socioeconomic group, smoking, and physical activity
Parity (mean)	Number of women			
1	270	reference	reference	reference
2	1,198	-0.26 (-0.86 to 0.33)	-0.07 (-0.68 to 0.53)	0.01 (-0.58 to 0.61)
3	584	0.07 (-0.58 to 0.73)	0.36 (-0.31 to 1.04)	0.30 (-0.36 to 0.96)
≥ 4 (4.4)	206	0.74 (-0.08 to 1.56)	1.08 (0.23 to 1.92)	0.81 (-0.02 to 1.63)
Total duration of breastfeeding (mean in months)				
Did not breastfeed	601	reference	reference	reference
<6 months (2.3)	752	-0.48 (-0.97 to 0.01)	-0.52 (-0.03 to -1.01)	-0.35 (-0.83 to 0.13)
6 - 9 months (7.3)	302	-0.53 (-1.16 to 0.10)	-0.55 (0.08 to -1.19)	-0.41 (-1.03 to 0.22)
≥ 10 months (18.5)	603	-0.75 (-0.24 to -1.26)	-0.93 (-0.39 to -1.46)	-0.50 (-1.03 to 0.03)
Change in mean BMI per 6 months breastfeeding*		-0.12 (0.004 to -0.25)	-0.18 (-0.05 to -0.32)	-0.10 (0.04 to -0.23)

* trend fitted using continuous duration of breastfeeding data (per month) and then multiplied by six to get the trend per six months

TABLE 4.5: Change in standardised mean BMI (kg/m^2) among postmenopausal women by parity and total duration of breastfeeding using BMI data from women with measured heights and weights and applying it to the entire study population using the regression dilution approach [30]

Parity (mean)	Number of women	Model A - standardised by age and region only			Model A + additionally standardised by breastfeeding (when looking at parity), and for parity (when looking at breastfeeding)		Model A + additionally standardised by breastfeeding, parity, socioeconomic group, smoking, and physical activity	
		reference	reference	reference	reference	reference	reference	reference
1	100,639							
2	310,841	0.01 (- 0.02 to 0.04)			0.12 (0.08 to 0.15)		0.13 (0.10 to 0.16)	
3	159,100	0.44 (0.41 to 0.48)			0.62 (0.58 to 0.66)		0.57 (0.54 to 0.61)	
≥ 4 (4.4)	80,600	1.26 (1.22 to 1.30)			1.48 (1.44 to 1.52)		1.29 (1.25 to 1.33)	
Total duration of breastfeeding (mean in months)								
Did not breastfeed	201,688		reference		reference		reference	
<6 months (2.3)	239,836	- 0.25 (- 0.22 to - 0.27)			- 0.30 (- 0.28 to - 0.33)		- 0.22 (- 0.20 to - 0.25)	
6 - 9 months (7.3)	82,198	- 0.39 (- 0.35 to - 0.42)			- 0.48 (- 0.45 to - 0.52)		- 0.33 (- 0.30 to - 0.37)	
≥ 10 months (18.5)	127,458	- 0.41 (- 0.38 to - 0.44)			- 0.71 (- 0.68 to - 0.75)		- 0.51 (- 0.49 to - 0.54)	
Change in mean BMI per 6 months breastfeeding*		- 0.05 (- 0.04 to - 0.06)			- 0.14 (- 0.13 to - 0.15)		- 0.11 (- 0.10 to - 0.12)	

* trend fitted using continuous duration of breastfeeding data (per month) and then multiplied by six to get the trend per six months

Bibliography

- [1] Bobrow K, Quigley M, Green J, Reeves G, Beral V. The long term effects of childbearing and breastfeeding on body mass index in middle aged women results from the Million Women Study. *Journal of Epidemiology & Community Health*. 2009;63(Suppl 2):56–56.
- [2] Whitlock G, Lewington S, Sherliker P, Clarke R, Emberson J, Halsey J, et al. Body-mass index and cause-specific mortality in 900 000 adults: collaborative analyses of 57 prospective studies. *Lancet*. 2009;373(9669):1083–1096.
- [3] Reeves GK, Pirie K, Beral V, Green J, Spencer E, Bull D, et al. Cancer incidence and mortality in relation to body mass index in the Million Women Study: cohort study. *BMJ*. 2007 Dec;335(7630):1134.
- [4] Sheldon JH. Maternal obesity. *Lancet*. 1949 Nov;2(6585):869–873.
- [5] Harris HE, Ellison GT, Holliday M. Is there an independent association between parity and maternal weight gain? *Annals of Human Biology*. 1997 Dec;24(6):507–519.
- [6] Gunderson EP, Abrams B. Epidemiology of Gestational Weight Gain and Body Weight Changes After Pregnancy. *Epidemiologic Reviews*. 1999 Jan;21(2):261–275.
- [7] Ohlin A, Rossner S. Factors related to body weight changes during and after pregnancy: the Stockholm Pregnancy and Weight Development Study. *Obesity research*. 1996 May;4(3):271–276.
- [8] Sichieri R, Field AE, Rich-Edwards J, Willett WC. Prospective assessment of exclusive breastfeeding in relation to weight change in women. *International Journal of Obesity & Related Metabolic Disorders: Journal of the International Association for the Study of Obesity*. 2003 Jul;27(7):815–820.
- [9] Heliövaara M, Aromaa A. Parity and obesity. *J Epidemiol Community Health*. 1981 Sep;35(3):197–199.
- [10] Kaye SA, Folsom AR, Prineas RJ, Potter JD, Gapstur SM. The association of body fat distribution with lifestyle and reproductive factors in a population study of postmenopausal women. *International Journal of Obesity*. 1990 Jul;14(7):583–591.
- [11] Bjorkelund C, Lissner L, Andersson S, Lapidus L, Bengtsson C. Reproductive history in relation to relative weight and fat distribution. *International Journal of Obesity & Related Metabolic Disorders: Journal of the International Association for the Study of Obesity*. 1996 Mar;20(3):213–219.

- [12] Weng HH, Bastian LA, Taylor DH, Moser BK, Ostbye T. Number of children associated with obesity in middle-aged women and men: results from the health and retirement study. *Journal of Women's Health* (2002). 2004 Feb;13(1):85–91.
- [13] Newby PK, Dickman PW, Adami H, Wolk A. Early anthropometric measures and reproductive factors as predictors of body mass index and obesity among older women. *Int J Obes Relat Metab Disord*. 2005;29(9):1084–1092.
- [14] Bastian LA, West NA, Corcoran C, Munger RG. Number of children and the risk of obesity in older women. *Preventive Medicine*. 2005;40(1):99–104.
- [15] World Health Organisation W. Global Database on Body Mass Index [online]; 2010. Internet [cited 02 November 2010]. Available from: http://apps.who.int/bmi/index.jsp?introPage=intro_3.html.
- [16] Ram KT, Bobby P, Hailpern SM, Lo JC, Schocken M, Skurnick J, et al. Duration of lactation is associated with lower prevalence of the metabolic syndrome in midlife–SWAN, the study of women's health across the nation. *American Journal of Obstetrics & Gynecology*. 2008 Mar;198(3):268.e1–268.e6.
- [17] Rush D, Lumey LH, Ravelli AC, Myers B. The indirect association of lactation with subsequent perimenopausal body weight. *European Journal of Clinical Nutrition*. 1996;50(1):12–16.
- [18] Hu F. *Obesity epidemiology*. Oxford: Oxford University Press US; 2008.
- [19] Wardle J, Waller J, Jarvis MJ. Sex Differences in the Association of Socioeconomic Status With Obesity. *Am J Public Health*. 2002 Aug;92(8):1299–1304.
- [20] Mujahid MS, Roux AVD, Borrell LN, Nieto FJ. Cross-Sectional and Longitudinal Associations of BMI with Socioeconomic Characteristics. *Obesity*. 2005;13(8):1412–1421.
- [21] Wamala SP, Wolk A, Orth-Gomér K. Determinants of Obesity in Relation to Socioeconomic Status among Middle-Aged Swedish Women. *Preventive Medicine*. 1997;26(5):734–744.
- [22] Molarius A, Seidell JC, Kuulasmaa K, Dobson AJ, Sans S. Smoking and relative body weight: an international perspective from the WHO MONICA Project. *Journal of Epidemiology and Community Health*. 1997 Jun;51(3):252–260.
- [23] Feskanich D, Willett W, Colditz G. Walking and Leisure-Time Activity and Risk of Hip Fracture in Postmenopausal Women. *JAMA*. 2002 Nov;288(18):2300–2306.
- [24] Tucker LA, Peterson TR. Objectively Measured Intensity of Physical Activity and Adiposity in Middle-Aged Women. *Obesity*. 2003 Dec;11(12):1581–1587.
- [25] The Million Women Study Collaborative Group. The Million Women Study: design and characteristics of the study population. *Breast Cancer Research: BCR*. 1999;1(1):73–80.
- [26] Townsend P, Phillimore P, Beattie A. *Health and deprivation*. Routledge; 1988.
- [27] Altman DG, Bland JM. Measurement in Medicine: The Analysis of Method Comparison Studies. *Journal of the Royal Statistical Society Series D (The Statistician)*. 1983 Sep;32(3):307–317.

- [28] Kirkwood BR, Sterne JAC. *Essential medical statistics*. Wiley-Blackwell; 2003.
- [29] Spencer EA, Appleby PN, Davey GK, Key TJ. Validity of Self-Reported Height and Weight in 4808 EPIC Oxford Participants. *Public Health Nutrition*. 2002;5(04):561–565.
- [30] Fletcher C, Peto R, Speizer F, Tinker C. *A follow-up study of the natural history of obstructive bronchitis*. Oxford University Press; 1976.
- [31] MacMahon S, Peto R, Collins R, Godwin J, MacMahon S, Cutler J, et al. Blood pressure, stroke, and coronary heart disease : Part 1, prolonged differences in blood pressure: prospective observational studies corrected for the regression dilution bias. *The Lancet*. 1990 Mar;335(8692):765–774.
- [32] Buka SL, Goldstein JM, Seidman LJ, Tsuang MT. Maternal Recall of Pregnancy History: Accuracy and Bias in Schizophrenia Research. *Schizophr Bull*. 2000 Jan;26(2):335–350.
- [33] Tomeo CA, Rich-Edwards JW, Michels KB, Berkey CS, Hunter DJ, Frazier AL, et al. Reproducibility and Validity of Maternal Recall of Pregnancy-Related Events. *Epidemiology*. 1999 Nov;10(6):774–777.
- [34] Promislow JHE, Gladen BC, Sandler DP. Maternal Recall of Breastfeeding Duration by Elderly Women. *Am J Epidemiol*. 2005 Feb;161(3):289–296.
- [35] Ko K, Park S, Kim Y, Bae J, Jun J, Gwack J, et al. Reliability of a questionnaire for women's reproductive history. *Journal of Preventive Medicine and Public Health Korean*. 2008;41(3).
- [36] Spencer EA, Roddam AW, Key TJ. Accuracy of self-reported waist and hip measurements in 4492 EPIC-Oxford participants. *Public Health Nutrition*. 2004 Sep;7(6):723–727.
- [37] Whitlock G, Clark T, Hoorn SV, Rodgers A, Jackson R, Norton R, et al. Random errors in the measurement of 10 cardiovascular risk factors. *European Journal of Epidemiology*. 2001;17(10):907–909.
- [38] Dugdale AE, Eaton-Evans J. The effect of lactation and other factors on post-partum changes in body-weight and triceps skinfold thickness. *British Journal of Nutrition*. 1989 Mar;61(2):149–153.
- [39] Lassek WD, Gaulin SJ. Changes in body fat distribution in relation to parity in American women: a covert form of maternal depletion. *American Journal of Physical Anthropology*. 2006 Oct;131(2):295–302.
- [40] Kjos SL, Henry O, Lee RM, Buchanan TA, Mishell DR. The effect of lactation on glucose and lipid metabolism in women with recent gestational diabetes. *Obstetrics and Gynecology*. 1993 Sep;82(3):451–455.
- [41] Hanley AJG, McKeown-Eyssen G, Harris SB, Hegele RA, Wolever TMS, Kwan J, et al. Association of parity with risk of type 2 diabetes and related metabolic disorders. *Diabetes Care*. 2002 Apr;25(4):690–695.
- [42] Blyton DM, Sullivan CE, Edwards N. Lactation is associated with an increase in slow-wave sleep in women. *Journal of sleep research*. 2002 Dec;11(4):297–303.
- [43] Mezzacappa ES, Kelsey RM, Katkin ES. Breast feeding, bottle feeding, and maternal autonomic responses to stress. *Journal of Psychosomatic Research*. 2005 Apr;58(4):351–365.

- [44] Bentley-Lewis R, Goldfine AB, Green DE, Seely EW. Lactation after normal pregnancy is not associated with blood glucose fluctuations. *Diabetes Care*. 2007 Nov;30(11):2792–2793.
- [45] Jonas W, Nissen E, Ransjö-Arvidson A, Wiklund I, Henriksson P, Uvnäs-Moberg K. Short- and long-term decrease of blood pressure in women during breastfeeding. *Breastfeeding Medicine: The Official Journal of the Academy of Breastfeeding Medicine*. 2008 Jun;3(2):103–109.
- [46] Ryan EA. Hormones and insulin resistance during pregnancy. *The Lancet*. 2003 Nov;362(9398):1777–1778.
- [47] Gunderson EP, Abrams B, Selvin S. The relative importance of gestational gain and maternal characteristics associated with the risk of becoming overweight after pregnancy. *International Journal of Obesity and Related Metabolic Disorders: Journal of the International Association for the Study of Obesity*. 2000 Dec;24(12):1660–1668.
- [48] Lawlor DA, Emberson JR, Ebrahim S, Whincup PH, Wannamethee SG, Walker M, et al. Is the Association Between Parity and Coronary Heart Disease Due to Biological Effects of Pregnancy or Adverse Lifestyle Risk Factors Associated With Child-Rearing?: Findings From the British Women's Heart and Health Study and the British Regional Heart Study. *Circulation*. 2003;107(9):1260–1264.
- [49] Stuebe AM, Rich-Edwards JW. The reset hypothesis: lactation and maternal metabolism. *American Journal of Perinatology*. 2009;26(1):81–88.

Chapter 5

Effects of childbearing on women's risk of diabetes after the menopause

Abstract

Objectives: *To explore the association between childbearing patterns and women's risk of diabetes after the menopause*

Design: *Cohort analysis of the risk of diabetes associated with women's childbearing histories*

Setting: *Population-based study of UK women recruited in 1996 to 2001 through the National Health Service Breast Screening Programme (the Million Women Study)*

Participants: *531,201 postmenopausal women who had not previously been diagnosed with diabetes and who reported their reproductive and medical histories, and other personal and sociodemographic factors*

Main outcome measure: *Risk of reporting a diagnosis of diabetes at first repeat survey and odds ratio adjusted for age, region, socioeconomic group, body mass index, smoking, alcohol, and strenuous physical activity*

Results: *Women were aged 57.6 years [SD 4] on average, were of mostly white ethnicity (98%), and had a mean BMI of 25.9 kg/m² [SD 4]. Of the 531,201 women included in the main analyses 7939 reported diabetes at repeat survey (1.5%) three years later. 88% of women were parous (mean parity 2.1 [SD 1.2]) and 71% of parous women reported ever breastfeeding (average total duration of breastfeeding 4.9 months [SD 7.9]). The overall risk of diabetes was 1.49%, the risk ranged from 1.40% for nulliparous women to 2.50% for women with parity four or more. Parous women who ever breastfed had a lower risk than parous women who never breastfed (1.43% as compared to 1.70%). In the unadjusted analyses the effects of parity and breastfeeding on women's risk of diabetes were small. Of the selected co-factors, body mass index (BMI, kg/m²) was most strongly associated with the risk of diabetes. After adjusting for age, region, socioeconomic group, BMI, smoking, alcohol, strenuous physical activity and breastfeeding there was no difference in the risk of diabetes between nulliparous and parous women (OR 1.05, CI 0.96 to 1.14). Among parous women, women of parity three or more had a slightly increased risk of diabetes compared to women with parity of one or two (OR 1.20, CI 1.08 to 1.34). Women who reported ever breastfeeding had a lower risk of diabetes (OR 0.89, CI 0.84 to 0.94) but there was no evidence to suggest a trend with increasing durations of breastfeeding ($p = 0.96$ for linear trend).*

Conclusions: *In this analysis of half-a-million postmenopausal women and almost 8000 cases of diabetes I found little evidence to suggest that parity or breastfeeding have a direct effect on women's risk of diabetes once their reproductive years are complete. The significant association between parity and breastfeeding and women's risk of diabetes was reduced after adjustment for BMI, which suggests that any association is likely to be mediated through BMI.*

5.1 Introduction

Diabetes is a common chronic disease that is responsible for significant morbidity and premature mortality. [1, 2] The incidence of diabetes in women increases with age and in the UK both the incidence and prevalence of diabetes in women have increased over time. [3–5] The risk of diabetes mellitus type II is known to be strongly affected by body mass index (BMI, kg/m^2), sociodemographic and lifestyle factors. [6–12] Evidence exists that the onset of the disease can be delayed through modification of these factors, for example maintaining a BMI in the “healthy” range (BMI 20 to 25 kg/m^2) and engaging in regular physical activity. [5, 13–15]

It has been suggested that women's childbearing histories may also influence their risk of diabetes in later life, however the existing evidence is unclear. We have already seen that there is good evidence that childbearing affects women's BMI after the menopause (see chapter 4, page 87). It is likely therefore that childbearing may, through its effects on BMI, affect women's risk of diabetes after the menopause. It is less certain that childbearing has any additional effects on women's risk of diabetes after the menopause.

For example, among populations of women who have completed their reproductive years, some investigators have reported finding an increased risk of diabetes with increasing parity, some report no association, and at least one investigator has reported a decreased risk of diabetes among parous women as compared to nulliparous women. [16–21] Where it has been possible to adjust for various measures of adiposity (most often BMI) investigators report a much weaker association between parity and maternal risk of diabetes. [17, 19, 21] Any residual increase in risk was limited to women in the highest parity categories (small heterogeneous groups of women likely to be different to the rest of the study population). In terms of breastfeeding, several investigators

have reported a decrease in risk of diabetes once their reproductive years are complete for women who have ever breastfed compared to those who never breastfed but few have found evidence of a trend associated with increasing duration. [21–24] Again, findings from these studies suggest that most of the effect was mediated through BMI. [21, 22, 24]

To date, the studies exploring the association between women's childbearing histories and their risk of diabetes in later life have been small and have focused on the independent effects of either parity or breastfeeding on risk. Only one study has explored the joint effects of parity and breastfeeding. [21] None of the studies have compared the relative size of the effects of childbearing as compared to known risk factors for diabetes.

In this chapter I report on the risk of reporting diabetes at first repeat survey within a large cohort of postmenopausal UK women in relation to their childbearing histories accounting for age, region, socioeconomic group, body mass index, smoking, alcohol, and strenuous physical activity. I also report on size of the effects of childbearing in relation to those of known important risk factors for diabetes.

5.2 Methods

5.2.1 Study Population

Recruitment

The Million Women Study is a population-based cohort of middle-aged women recruited via the National Health Service (NHS) Breast Screening Programme in England

and Scotland between 1996 and 2001. A total of 1.3 million women completed the first study questionnaire which asked about their reproductive and medical histories, including whether they had ever been diagnosed with or were being treated for diabetes as well as personal, sociodemographic, and lifestyle characteristics.

Repeat survey

Starting in 1999 the first round of repeat surveys were mailed to study participants about 3 years after they were recruited (65% response rate). In addition to updating lifestyle and other characteristics, women were again asked about their medical histories including whether they had been diagnosed with diabetes in the preceding 5 years (if yes, when) and their current medications.

Full details of the study design and methods have been published elsewhere, [25] aspects of the study design and methods relevant to this thesis are described in chapter 2 (page 18). Study questionnaires can be viewed at <http://www.millionwomen.org> and have been included in this thesis in Appendix A, page 241. The study was approved by the Anglia and Oxford Multi-Centre Research Ethics Committee. All participants gave their written consent to take part in the study.

5.2.2 Data Definitions

Menopause

Women were classified as postmenopausal if they reported their periods stopped naturally or after bilateral oophorectomy, or if they were aged 53 years or older (most

women aged 53 years who had a natural menopause were postmenopausal by that age).

Parity and breastfeeding

At recruitment, women were asked to report the number of children they had given birth to (including still births), which was used to define parity. Women were asked to report, for each birth, if they had breastfed and if so, the duration of breastfeeding in months. This information was used this information to define breastfeeding (ever or never), and total duration of breastfeeding (summation over all children of reported duration of breastfeeding in months). In addition to using women's total duration of breastfeeding I also calculated the average duration of breastfeeding per child using each woman's total duration of breastfeeding divided by her parity. No questions were asked about the exclusivity of breastfeeding.

Diabetes

At recruitment women were asked "Have you ever had [Diabetes]?" and "Are you now being treated for [Diabetes]?" Women who answered yes to either of these questions were considered to have prevalent diabetes. At the first repeat survey around 3 years, later women were asked "Have you had any of the following conditions [Diabetes] diagnosed for the first time in the last 5 years?", and "Are you now being treated for: [Diabetes] If yes, about how old were you when treatment began?" Women who answered yes to either of these questions were considered to have diabetes mellitus reported at

resurvey. No questions were asked about a history of diabetes in pregnancy. The outcome of interest analysed in this chapter was diabetes reported at repeat survey among women who had not previously been diagnosed with diabetes.

The reliability and validity of self-reported diabetes among Million Women Study participants have been discussed previously (see chapter 2, section 2.3.3, page 25). Briefly for this study population, self-reported diabetes has been shown to be valid measure (98% agreement between women's self-report of treatment for diabetes and general practitioner records for treatment of diabetes, kappa 0.78). [26] In addition 83% of women who reported diabetes at recruitment also reported diabetes at the first repeat survey (table 2.6, page 57). These findings agree with existing literature. [27–33] Although it was not possible to classify women by sub-type of diabetes (type I or type II) given the mean age of the study population (56 yrs), mean age at diagnosis (60 yrs) and the high percentage of overweight and obesity (53%), it is likely that most of the new onset diabetes cases are due to diabetes mellitus, type II. [2]

Co-factors & effect modifiers

Using data from the recruitment questionnaire participants in the Million Women Study were classified into 10 geographic regions of residence within England and Scotland. Women's socioeconomic status was derived using the Townsend deprivation index and summarised into quintiles. [34] Body mass index (BMI) was calculated as a woman's weight in kilograms divided by her height in meters squared. Adiposity as captured by BMI is a well established risk factor for diabetes. [6, 7, 35, 36] Where appropriate I summarised women's BMI into fine categories (<20, 20 to 22.4, 22.5 to 24.9, 25 to 27.4, 27.5 to 29.9, 30 to 32.4, 32.5 to 35, or $\geq 35 \text{ kg/m}^2$). I used 22.5 to 24.9 kg/m^2 as the reference category. I also presented results of the association between childbearing and

diabetes by category of BMI, using the following BMI categories <25, 25 to 29.9, 30 to 34.9, or ≥ 35 kg/m². Women were categorised as never, past, or current smokers. Current smokers were further categorised by the number of cigarettes smoked per day as follows, 1 to 4, 5 to 9, 10 to 14, 15 to 19, 20 to 24, or ≥ 25 . Women's alcohol intake was summarised using information on the number of glasses of wine, half pints of beer/lager, or measures of spirits drunk in an average week. This information was converted into standard number of drinks consumed per week and alcohol intake was categorised as none, 1 to 6, 7 to 14, or ≥ 15 units per week. Strenuous physical activity was summarised into four categories, none, ≤ 1 time per week, 2 to 3 times per week, 4 or more times per week. Information on co-factors was taken as reported at recruitment. Where appropriate I adjusted for measurement error in smoking, and alcohol using the regression dilution approach, I summarised each of the co-factors into categories based on information from the recruitment questionnaire, however the average for each category was calculated for smoking and alcohol using data as reported on the repeat survey questionnaire (administered about 3 years after recruitment) from 708,265 women. [37]

5.2.3 Statistical Analysis

Eligibility and exclusion criteria

Women were eligible for inclusion into the main analysis if they were postmenopausal, did not have a registered diagnosis of cancer (except non-melanoma skin cancer) at recruitment, had been asked about their breastfeeding histories, and had completed and returned a first repeat survey questionnaire. Of the 598,803 women who were eligible,

women missing information on parity (0.13%), breastfeeding (9%), or diabetes at recruitment (0.02%) were excluded. For these analyses, women with prevalent diabetes at recruitment were also excluded (2.24%). The final number of women included in the main analysis was 531,201 (88% of those eligible). Women missing values for any of the adjustment variables were assigned to a separate category for that variable. Figure 5.1 (page 145) shows a flow diagram of the eligibility and inclusion criteria and the numbers of women excluded.

Exposures, outcome, co-factors, & effect modifiers

The main exposure variables were parity and duration of breastfeeding treated as categorical variables. Parity was summarised as 0, 1, 2, 3, or 4 or more; and total duration of breastfeeding as never breastfed, breastfed for less than 6 months, breastfed for 6 to 9 months, or breastfed for 10 months or more (these approximate tertiles of total duration for women who breastfed). The outcome of interest was risk of diabetes at repeat survey. All analyses were adjusted for age (5 year categories), socioeconomic group (in fifths), and region of residence (10 regions). Univariate analyses were performed to assess the following co-factors identified a-priori - smoking, alcohol, and strenuous physical activity. Co-factors found to be significantly associated with women's risk of diabetes at repeat survey ($p < 0.0001$) were included in multivariate analyses. BMI was considered to be an important potential effect mediator a-priori, associations between childbearing and diabetes were variously modeled to explore for this.

Model specifics

The association between risk of reporting diabetes at repeat survey and childbearing history was assessed using logistic regression models. Women of parity two (mean parity in this study population was 2.1 children [SD 1.2]) who never breastfed were used as the reference group in these analyses. I calculated crude risks and as discussed in chapter 2 (section 2.5, page 47) I applied floating absolute risk methods to calculate floated odds. [38] Where appropriate I used the category means to performed trend tests.

Additional analyses

Excess adiposity as captured by BMI (kg/m^2) is an important known risk factor for diabetes. [6, 7, 13, 35, 36, 39] Work presented in chapter 4, page 87 showed clear evidence that both parity and breastfeeding affect women's BMI after the menopause. (See figure 4.7, page 115). As it is possible that any effects of childbearing on women's subsequent risk of diabetes are mediated through their effects on women's adiposity (in other words BMI is part of the causal pathway), I examined the independent effects of parity and breastfeeding in models with and without adjusting for BMI. I also examined the effects of parity and breastfeeding within strata of BMI. To assess the potential effects of measurement error on model estimates due to the use of self-reported BMI, I repeated the main analysis using the regression dilution approach. [37, 40] BMI was summarised into fine categories as described above (using self-reported data at recruitment), however the average for each category of BMI was calculated using measured height and weight data taken from a randomly selected sample of women ($n=3745$) eight years after recruitment. [41] I repeated the main analysis using this data.

I carried out additional analyses, using information obtained at repeat survey, looking at the potential effects of women's reported ethnicity, a family history of diabetes, and her own birth weight (in kilograms) classified as ≤ 2.4 kg, 2.5-3.4 kg, 3.5-4.4 kg, or >4.5 kg on model estimates. I also performed a sensitivity analysis restricting the main analysis to women with complete information for all relevant factors. All analyses were performed using STATA version 11 (State corporation, TX, USA).

5.3 Results

Characteristics of the study population

Table 5.1 (page 146) shows the baseline characteristics of women included in this analysis. The risk of diabetes in this population was 1.5% (531,201 women were included of whom 7,939 developed diabetes). Women were aged 57.6 years on average (SD 4) and were mostly of white ethnicity (98%). Most women were parous (88%) with a mean parity of 2.1 children (SD 1.2) and average age at first birth of 24.1 years (SD 4.2). 71% of parous women reported ever breastfeeding; the mean number of children breastfed was 1.2 (SD 1.2) and the average duration of breastfeeding per child was 3.2 months (SD 3.1).

It can be seen from table 5.1 (page 146) that nulliparous women were less deprived, had lower BMIs and were less likely to be overweight, less likely to smoke, consumed more alcohol, and reported engaging in strenuous physical activity more often than parous women. Among parous women, mean BMI and the proportion of women with a mean BMI ≥ 25 increased with increasing parity (58% of women with parity of four or more) as did the risk of reporting diabetes at repeat survey (women of parity two had the

lowest risk, 1.2%). The proportion of parous women who reported being current smokers also increased with increasing parity. Average intake of alcohol and the amount of strenuous physical activity women reported were similar for all parity groups. In terms of breastfeeding histories, the proportion of women reporting ever breastfeeding, the number of children breastfed and the total duration of breastfeeding all increased with increasing parity. Women who reported ever breastfeeding were less deprived, had lower BMIs, were much less likely to be current smokers, and were more likely to report regular strenuous physical activity than women who never breastfed. There was no difference by parity of breastfeeding status in women's self-report of a family history of diabetes (12%).

Table 5.2 (page 147) shows the differences in selected recruitment characteristics for women who did and who did not complete and return the first repeat survey questionnaire. Questionnaire response rates were higher for women from regions in the south of England, of higher socioeconomic status, with a lower BMI, who regularly consumed alcohol, and who never smoked. Comparing women who did and who did not complete and return the first repeat survey questionnaire, there was no difference in mean parity. Women who did not return a questionnaire were 10% less likely to have ever breastfed. Among parous women who reported having ever breastfed there was very little difference in the number of children or duration of breastfeeding between women who did and did not complete the repeat survey.

Univariate analyses

Table 5.3 (page 148) shows the risk of diabetes for selected characteristics in models adjusted for age, and region. Compared to parous women, a history of not having

had any children was associated with a moderate decrease in the risk of diabetes. For parous women having ever breastfed was associated with a modest decrease in the risk of diabetes. The co-factor associated with the biggest effect on diabetes risk was BMI; the risk of diabetes associated with a BMI of more than 30 kg/m² was at least an order of magnitude larger than any of the other risk factors considered.

Figures 5.2 - 5.6 (pages 149 - 151) show the odds ratios (95% CI) of reporting diabetes by parity, duration of breastfeeding (limited to parous women), smoking, BMI, and levels of strenuous physical activity. Except for BMI, the figures are all on a similar scale. Each model was adjusted for age, region, and socioeconomic group.

From these figures it can be seen that the small effects on women's risk of diabetes associated with childbearing were similar to the size of effects of smoking and strenuous physical activity. The scale of these figures is the same. In contrast, the scale of the figure showing the association between BMI and women's risk of diabetes is four times greater. It can be seen in figure 5.4 (page 150) that BMI was associated with a large and exponential increase in women's risk of diabetes. For example, women with a BMI of 35 kg/m² or more were 16 times more likely to develop diabetes than women with a BMI of 22.5 to 24.9 kg/m².

Multivariate analyses

Table 5.4 (page 152) shows the odds ratios for diabetes by parity and by breastfeeding status in models with various degrees of adjustment. Results shown in the table are discussed below.

Comparing nulliparous to parous women

In a model (Model A) adjusted only for age, region, and socioeconomic group being parous was associated with a moderate increase in the risk of diabetes. Adjusting for BMI weakened this association. In the fully adjusted model, with adjustment for age, region, socioeconomic group, BMI, smoking, alcohol, and strenuous physical activity, there was no significant difference in the risk of diabetes between nulliparous and parous women (OR 1.05, CI 0.96 to 1.14).

Risk associated with increasing parity

In table 5.4 (page 152) results from Model A show that, for parous women, increasing parity was associated with an excess risk of diabetes. Again adjusting for BMI weakened this association. In the fully adjusted model (adjusted as described above) increasing parity was still associated with a moderate increase in the risk of diabetes. It is also clear from table 5.4 (page 152) that adjusting for adiposity as captured by BMI had the single biggest effect on the risk estimate between increasing parity and women's risk of diabetes.

Comparing women who breastfed to women who did not breastfeed

Compared to parous women who never breastfed, a history of having ever breastfed was associated with a lower risk of diabetes in Model A. This association was weakened after adjusting for BMI. In the fully adjusted model a history of having ever breastfed was still associated with a modest decrease in parous women's risk of diabetes.

Risk among parous women by duration of breastfeeding

Increasing durations of breastfeeding were associated with a decrease in the risk of diabetes among parous women in Model A. For example, compared to parous women who did not breastfeed, the odds ratio decreased to 0.70 (CI 0.61 to 0.80) for women who breastfed for ten months or more. Adding adjustment for BMI weakened this association. In the fully adjusted model the protective effect associated with breastfeeding was small, and similar across categories of breastfeeding duration. It is also notable that the risk estimates for a history of having ever breastfed or any of the categories of total duration of breastfeeding were very similar. For most of the models preceding the main model, tests for linear trend of the association between increasing durations of breastfeeding and women's risk of diabetes were borderline significant (p-value between 0.04 to 0.07). In the fully adjusted model the statistical evidence for a linear trend was again borderline (p=0.06). When a Likelihood Ratio-test for trend associated with increasing durations of breastfeeding was restricted to women who had breastfed there was no statistical evidence for a linear trend (p=0.96). When duration of breastfeeding (in all women) was modeled as a continuous variable there was no statistical evidence that increasing durations of breastfeeding were associated with the risk of diabetes in this population (p=0.11).

Findings in context**The contribution of BMI to women's risk of diabetes**

The findings presented in table 5.3 (page 148) showed that BMI was strongly associated with women's risk of diabetes.

It can be seen from table 5.4 (page 152) that adjusting for BMI weakened the association between childbearing and women's risk of diabetes. Results presented in the table also show that of the co-factors included in the fully adjusted model BMI had the biggest effect on the association between childbearing and diabetes risk.

Table 5.5 (page 153) shows the values of the χ^2 statistic for heterogeneity across categories of parity and also across categories of total duration of breastfeeding (on three degrees of freedom) for each of the models in table 5.4 (page 152). It illustrates the relative contributions to the value of the χ^2 statistic of adjusting for various co-factors. For both parity and breastfeeding adjusting for BMI had the biggest effect on the associated χ^2 statistic, reducing the associated χ^2 value by about two thirds.

The effect of stratifying by BMI

Table 5.7 (page 155) presents results of the risk of diabetes associated with childbearing by broad strata of BMI (<25 , 25 to 30 kg/m², 30 to 35 kg/m², and ≥ 35 kg/m²)¹. It can be seen that the association between childbearing and risk of diabetes wanes with increasing BMI suggesting that BMI may be mediating the association. Figures 5.7 (page 156) show that increasing BMI interacts with childbearing and women's risk of diabetes.

The contributions to risk of childbearing and selected cofactors

The results presented in table 5.6 show the effect size of the associations between childbearing and selected cofactors and women's risk of diabetes. The risk of diabetes associated with childbearing after full adjustment was modest in comparison to other

¹in a model adjusted for age, region, socioeconomic group, smoking, alcohol, strenuous physical activity, and parity and breastfeeding as appropriate

co-factors. The strong association between increasing BMI and risk of diabetes is clear. In a model which included BMI as a continuous variable for women with a BMI of 25 kg/m² or higher an increase of one additional unit of BMI was associated with a 14% increase in the risk of diabetes (OR 1.14, CI 1.13 to 1.14). These results can be clearly seen in the figure set (figures 5.8 - 5.13) on pages 157 - 160.

Figure 5.10 (page 159) shows the association between BMI and risk of diabetes. The plot shows the floated odds ratio and the y-axis scale on this figure is from 0 to 16. For comparison similar graphs showing the odds ratios of diabetes associated with parity (figure 5.8 , page 157), duration of breastfeeding (figure 5.9, page 158), socioeconomic group (figure 5.12, page 160), smoking (figure 5.11, page 159), and strenuous physical activity (figure 5.13, page 160) have also been plotted however the y-axis scale on these graphs is from 0 to 4. The strong association between increasing BMI and risk of diabetes is clear as is the size of the risks associated with BMI relative to the other factors in this analysis. The flatness of the graph for an effect of increasing duration of breastfeeding is also clear.

Additional analyses

Effects on model estimates of other cofactors which may influence women's risk of diabetes

Table 5.8 (page 161) shows the results of associations between childbearing and women's risk of diabetes from the fully adjusted model and from models with additional adjustment for either ethnicity, family history of diabetes, or mother's own birth-weight. It is clear from the table that none of these factors materially affect the overall results. For example, the risk of diabetes associated with a parity of three was 1.16 (CI 1.09 to 1.24)

in the fully adjusted model, in comparison the HR in models additionally adjusted for ethnicity, family history of diabetes, or mothers own birth weight the risk estimate for diabetes associated with a parity of three were 1.16 (CI 1.11 to 1.22), 1.16 (CI 1.11 to 1.22), and 1.16 (CI 1.10 to 1.22) respectively.

Effect on model estimates of using measured heights and weights to calculate BMI and the regression dilution technique to correct for misclassification

Table 5.9 (162) shows the the results of the associations between childbearing and women's risk of diabetes from two models which were similarly adjusted except for BMI. The first model (on the left) BMI was included in the model as described in the methods section. For the second model (on the right) BMI was included using the regression dilution technique. [37, 40] Categories of BMI were the same as in the main analysis (constructed using self-reported height and weight data taken at recruitment) however the category means were calculated using using measured heights and weights from a random sample of women (taken about 8 years after recruitment). There was no appreciable difference in the two models (differences in the models occurred at the fifth decimal point).

5.4 Discussion

Summary

In this analysis of over half-a-million women and nearly 8000 cases of diabetes I found evidence that women's childbearing histories were associated with their risk of developing diabetes after the menopause that was largely mediated through their effects on

women's BMI. I found limited evidence to suggest a direct effect of childbearing on women's risk of diabetes.

The residual risk of diabetes associated with increasing parity was statistically significant and similar in size to the risks of diabetes associated with socioeconomic group, smoking, and strenuous physical activity, however these were dwarfed by the risk associated with increasing BMI. For example, comparing the risk estimates for selected variables included in the final model parity four or more OR 1.36 (CI 1.28 to 1.47), breastfeeding 10 or more months OR 0.87 (CI 0.81 to 0.93), lowest socioeconomic group OR 1.17 (CI 1.11 to 1.23), current smoker (10 to 14 cigarettes per day) OR 1.54 (CI 1.37 to 1.72), no strenuous physical activity OR 1.25 (CI 1.22 to 1.29), with BMI 30 to 32.5 kg/m² OR 6.08 (CI 5.63 to 6.57). Although a history of having ever breastfed was associated with a small (statistically significant) decrease in the risk of diabetes after full adjustment, there was no evidence of a linear trend with increasing durations of breastfeeding.

Findings in relation to other studies

In terms of parity, the finding that after appropriate adjustment for BMI there was no difference in the risk of diabetes between nulliparous women and parous women has been reported previously. [21, 42] These findings agree with published evidence which has suggested that increasing parity is associated with a small increase in the risk of diabetes. [21, 43, 44] In keeping with other published data the biggest effect was seen in the highest parity category. [19] Women in this category are only a small proportion of the study population, and there is often a large variation within this category (for

example parity between 5 and 10 in our study population), and have different lifestyle characteristics to other study participants (see Table 5.1, page 146).

In terms of breastfeeding, investigators have consistently reported an association in parous women between reporting having ever breastfed and a lower risk of diabetes² [21, 24] Findings on the effects of duration of breastfeeding have been less consistent. As with the findings reported in three prospective cohort studies (the Nurses Health Study I, the Shanghai Women's Study, and the Reproductive Risk Factors for Incontinence Study at Kaiser) I did not find evidence of a dose-response trend with increasing durations of breastfeeding and risk of diabetes. [22–24] In contrast, published findings from the Nurses Health Study II and a cross-sectional analysis of the 45 and Up cohort in Australia reported a negative association between increasing durations of breastfeeding and risk of diabetes.[21, 22] Women from these two studies may be quite different to women included in our analysis. For example, in the 45 and Up cohort women had a mean parity of 2.8 (SD 1.3) and average duration of breastfeeding 15.9 months as compared to women in our study who had a mean parity of 2.1 (SD 1.2) and an average duration of breastfeeding of 4.9 months. (Dr Liu - personal communication)

Possible explanations

It is known that BMI is an imperfect measure of adiposity.[36] For example it does not capture the distribution of adipose tissue and the relative excess of adipose tissue at certain sites may be at least as important to the risk of disease as overall excess adiposity. [45, 46] In this analysis adjusting for BMI had a considerable effect on the model estimates and the size of the χ^2 statistic which underlines the strong effect of adiposity (however imperfectly measured) on the risk of diabetes. If the effect of breastfeeding

²although recently reports have focused on the excess risk associated with not breastfeeding

on women's risk of diabetes was primarily mediated through BMI then the largest effect would be noticed between women who did and did not breastfeed and it would be similar to the size of the effect of breastfeeding on BMI.

In this analysis, for women with BMI of 25 kg/m^2 , a one unit increase in BMI was associated with a 14% increase in a woman's risk of diabetes. In chapter 4 (page 87) it was seen that for parous women, a history of breastfeeding was associated with a standardised mean BMI roughly one third of a unit lower than otherwise similar women who did not breastfeed. Given this finding it is reasonable to expect that parous women who breastfed would have about a 5% lower risk of diabetes than women who never breastfed due to the difference in their BMIs after the menopause. In the analyses presented in this chapter the risk of diabetes associated with a history of breastfeeding was 11% (OR 0.89, CI 0.84 to 0.94) lower than women who did not breastfeed. It is notable that the 5% difference we might expect based on the earlier finding of the size of the effect on BMI associated with ever breastfeeding falls within the 95% confidence intervals of the risk estimate of diabetes associated with breastfeeding. It is also notable that there was no statistical evidence of a trend in risk decrease with longer durations of breastfeeding.

Excess adiposity as captured by BMI is well known to be an important risk factor for diabetes, [39, 47–49] I found a strong and independent association with increasing BMI and risk of developing diabetes that was an order of magnitude greater than other co-factors and the exposures of interest (if there is truly an independent association between parity, breastfeeding and risk of diabetes). The findings in regards to the risk of diabetes associated with socioeconomic group, physical activity, and alcohol are consistent with the published literature.[13, 50–53]

Since it is not possible to eliminate residual confounding due to incompletely measured adiposity it is prudent to interpret the direct effects of childbearing on risk of diabetes cautiously. It is plausible that the effects of childbearing on women's risk of diabetes are indirect and are mediated through their association with maternal BMI after the menopause. There is as yet very little evidence to suggest a biologically plausible mechanism by which childbearing has a meaningful effect on glucose homeostasis that is independent of BMI.

Strengths and weaknesses

The strengths of this study include its size and the availability of information about co-factors. The Million Women Study includes one in four UK women aged between 50 and 64 years at the time of recruitment. This analysis included 531 201 women and 7939 cases of diabetes which to my knowledge is the largest sample-size to date. As with most large-scale epidemiological studies exposure information was obtained by self-report. Studies in older women have found that their long-term recall of parity and breastfeeding is reliable, [54, 55] and I have shown previously that self-reported breastfeeding data corresponds well to data obtained from historical data sets (Chapter 3, page 63). Given the historical trends in breastfeeding in this population of women there is substantial power to assess the effects of short durations of breastfeeding but less power to assess the effects of long durations of breastfeeding if they exist. I used self-reported diagnosis of or treatment for diabetes as our outcome variable in this analysis. The reliability and validity of self-reported diabetes has been shown previously and echoes findings in the published literature. [26–33] I have detailed information on

socioeconomic factors, smoking and physical activity which allowed for fine subdividing by key co-factors. I also have updated exposure information which allowed me to correct for possible misclassification errors. [37, 40]

Conclusions

In this analysis of a large population of UK women I found evidence that both parity and breastfeeding were associated with the risk of developing diabetes after the menopause that was largely mediated through BMI(kg/m²). I found limited evidence to suggest a direct effect of childbearing on women's risk of diabetes.

5.5 Chapter summary

In this chapter I explored the associations between women's childbearing histories and their risk of developing diabetes after the menopause. Building on the findings presented in chapter 4 (page 87) which showed that childbearing was meaningfully associated with maternal BMI(kg/m²) after the menopause I explored the relative associations of childbearing, BMI(kg/m²), and other co-factors with women's risk of diabetes after the menopause. I found BMI was very strongly associated with women's risk of developing diabetes, a finding that is consistent with the published literature. I investigated how accounting for BMI through adjustment and through stratification influenced the associations between childbearing and women's risk of diabetes after the menopause. I presented evidence that both parity and breastfeeding were meaningfully associated with the risk of developing diabetes after the menopause but that this was largely mediated through BMI(kg/m²).

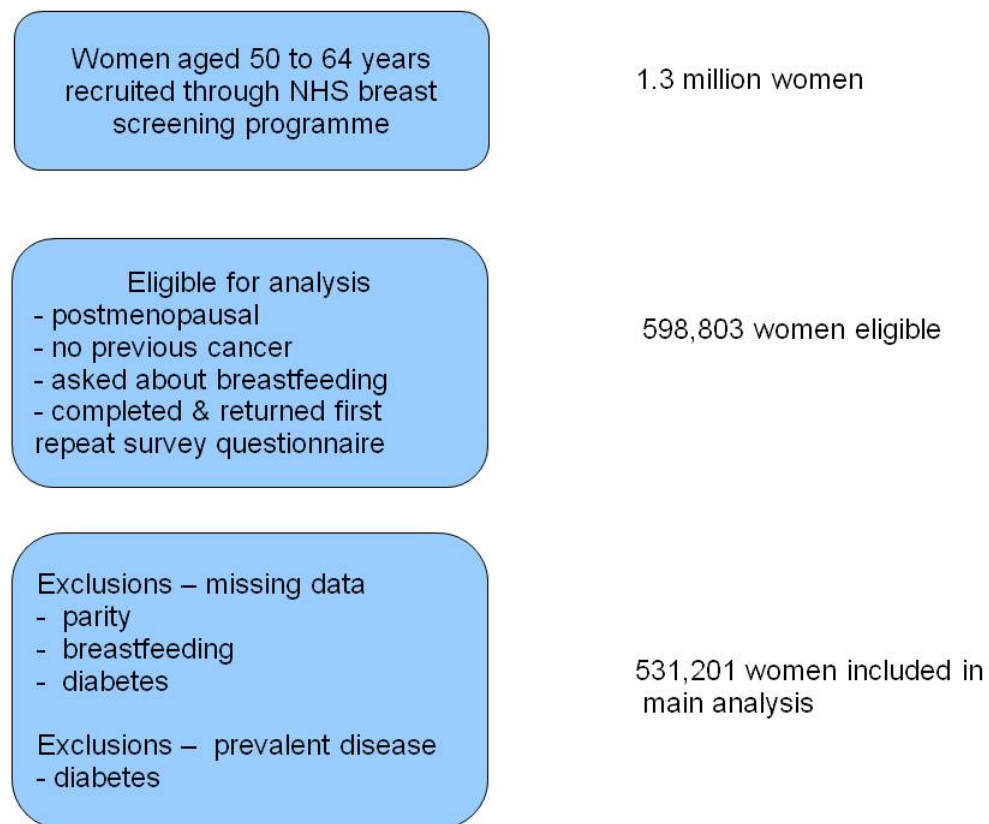


FIGURE 5.1: Flow diagram showing eligibility and inclusion criteria and the final number of women included in the main analyses

TABLE 5.1: Characteristics of the study population according to parity and breastfeeding status

Population Characteristics	By Parity				By Breastfeeding Status*		
	Nulliparous n = 65 090	One	Two	Three	Never Breastfed n = 132 611	Ever Breastfed n = 333 500	All Women n = 531 201
		n = 70 376	n = 229 285	n = 113 845			
Mean age in years (SD)	57.5 (5)	57.5 (5)	57.2 (4)	57.8 (4)	56.7 (4)	57.9 (4)	57.6 (4)
Mean parity (SD)	0	1	2	3	2.1 (0.9)	2.5 (1)	2.1 (1.2)
Breastfeeding							
Number of children breastfed (Mean [SD])	0	0.5 (0.5)	1.2 (0.8)	1.8 (1.2)	0	1.9 (1.0)	1.2 (1.2)
Total lifetime duration in months (SD)	0	2.0 (3.3)	4.5 (5.9)	7.4 (8.8)	0	7.8 (8.7)	4.9 (7.9)
Lower third of socioeconomic group (%)	33	34	30	35	37	32	33
Post-high school qualification (%)	23	14	16	15	9	17	16
Body mass index (Mean (kg/m2) [SD])	25.5 (5)	25.8 (4)	25.8 (4)	26.2 (4)	26.2 (5)	25.9 (4)	25.9 (4)
BMI 25kg/m ² or higher (%)	45	47	48	52	52	50	50
Current smokers (%)	13	16	13	15	18	13	15
Daily alcohol consumption (standard drinks/day)	1.1 (6.6)	1.0 (6.7)	1.0 (5.9)	1.0 (6.6)	1.0 (6.4)	1.0 (6.4)	1.0 (6.4)
Reporting engaging in strenuous physical activity more than once per week (%)	25	24	24	24	22	25	24
Total number of women who developed diabetes during follow-up	909	1019	2833	1862	2257	4773	7939
Risk of diabetes	1.40	1.45	1.24	1.64	1.70	1.43	1.49

* Amongst parous women only

Percentages exclude the small number of women with missing values

TABLE 5.2: Selected recruitment characteristics of women who did and who did not return a completed first repeat survey questionnaire

Selected characteristics at study recruitment	Women who completed the first repeat survey	Women who did not complete the first repeat survey
Number of women (%)	866,551 (63)	497,744 (36)
Age in years (Mean [SD])	56 [5]	56 [5]
Parity (Mean [SD])	2.1 [1.2]	2.2 [1.3]
Number of children breastfed (Mean [SD])*	1.2 [1.2]	1.1 [1.2]
Women who ever breastfed (%)*	62	55
Total duration of breastfeeding in months* (Mean [SD])	5 [8]	4 [8]
In lower third of socioeconomic group (%)	30	40
Current smokers (%)	16	28
Physical activity less than once a week (%)	55	44

* Women who were asked about breastfeeding

TABLE 5.3: Odds ratio of diabetes among postmenopausal women for selected characteristics in models adjusted only for age and region (OR, 95%CI)

Characteristics of women		Cases	Population at risk	Risk (%)	OR	(95% CI)
Parity						
	Nulliparous	909	65,090	1.40	1.00	(0.82 - 1.22)
	Parous	7,030	466,111	1.51	1.07	(0.87 - 1.33)
LR test heterogeneity					p = 0.042	
Breastfed						
	No	2,257	132,611	1.70	1.00	(0.82 - 1.22)
	Yes	4,773	333,500	1.43	0.87	(0.71 - 1.06)
LR test heterogeneity					p < 0.0001	
Socioeconomic group						
	Upper third	1,999	175,003	1.14	0.82	(0.68 - 1.00)
	Middle third	2,438	178,429	1.37	1.00	(0.83 - 1.20)
	Lowest third	3,502	177,769	1.97	1.42	(1.17 - 1.73)
LR test heterogeneity					p < 0.0001	
Body mass index						
	< 20	58	18,947	0.31	0.62	(0.48 - 0.80)
	20 to 22.4	208	83,046	0.25	0.51	(0.45 - 0.59)
	22.5 to 24.9	714	144,049	0.50	1.00	(0.94 - 1.07)
	25 to 27.4	1,169	113,091	1.03	2.06	(1.95 - 2.19)
	27.5 to 29.9	1,400	70,405	1.99	4.00	(3.79 - 4.22)
	30 to 32.4	1,238	37,749	3.28	6.71	(6.33 - 7.12)
	32.5 to 34.9	1,005	21,124	4.76	9.95	(9.33 - 10.6)
	≥ 35	1,663	21,429	7.76	17.05	(16.2 - 17.9)
LR test heterogeneity					p < 0.0001	
Smoking						
	Never	3,634	278,191	1.31	1.00	(0.91 - 1.09)
	Past	697	36,450	1.91	1.56	(1.39 - 1.75)
	Current <15	2,497	148,836	1.68	1.29	(1.17 - 1.43)
	Current >15	645	40,981	1.57	1.25	(1.11 - 1.40)
LR test heterogeneity					p < 0.0001	
Strenuous physical activity						
	< 1 per week	1,999	170,730	1.17	1.00	(0.91 - 1.10)
	≥ 1 per week	1,055	115,721	0.91	0.77	(0.61 - 0.97)
LR test heterogeneity					p < 0.0001	

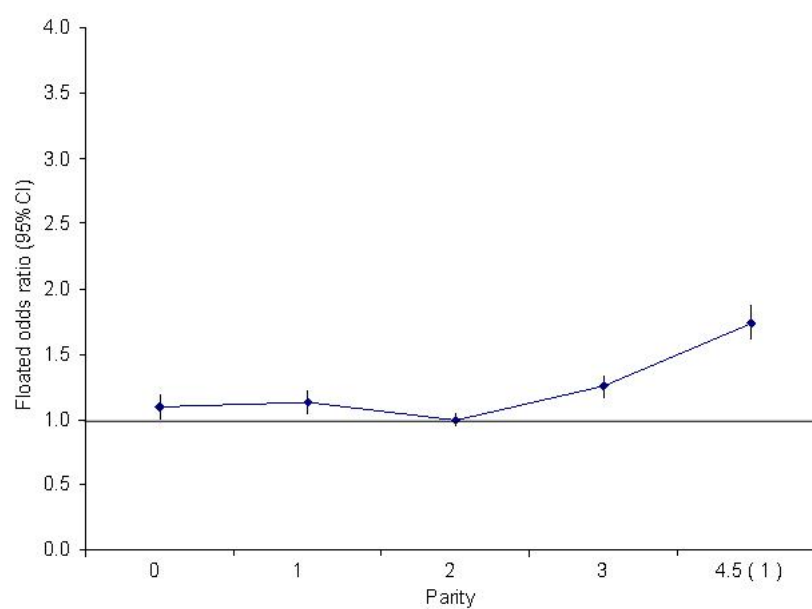


FIGURE 5.2: Odds ratio of diabetes by parity in a model adjusted for age, region, and socioeconomic group (OR, 95%CI)

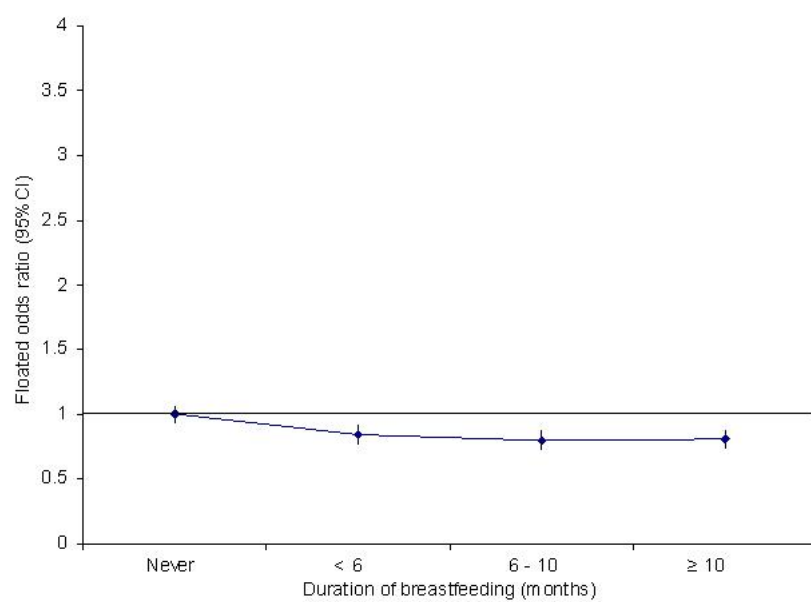


FIGURE 5.3: Odds ratio of diabetes among parous women by duration of breastfeeding in a model adjusted for age, region, and socioeconomic group (OR, 95%CI)

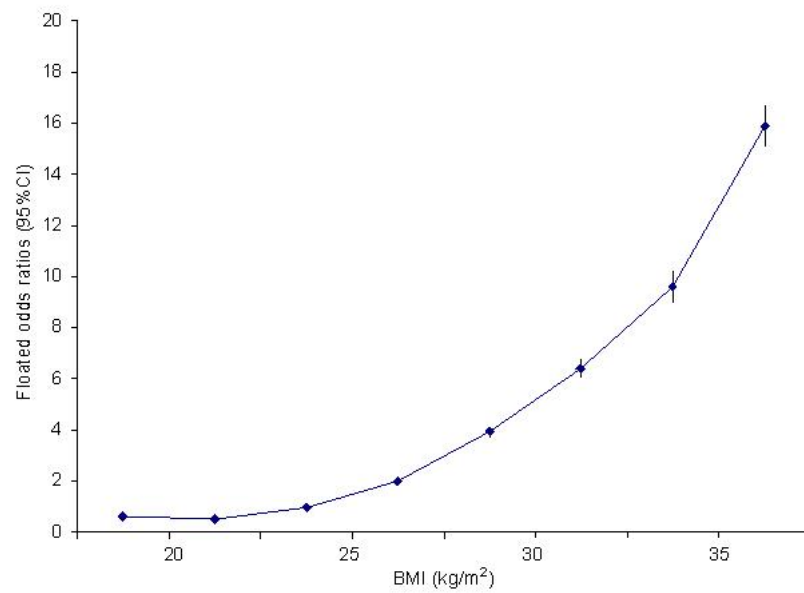


FIGURE 5.4: Odds ratio of diabetes by BMI in a model adjusted for age, region, and socioeconomic group (OR, 95%CI)

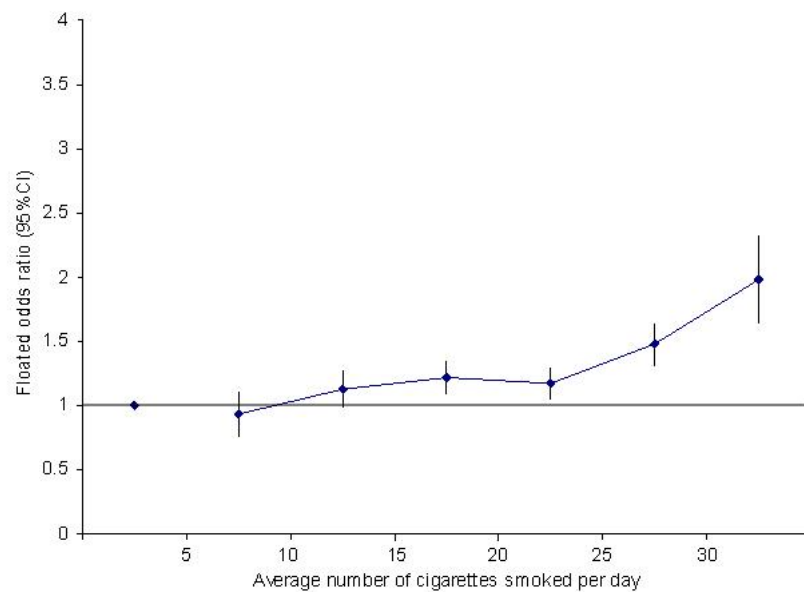


FIGURE 5.5: Odds ratio of diabetes by number of cigarettes smoked in a model adjusted for age, region, and socioeconomic group (OR, 95%CI)

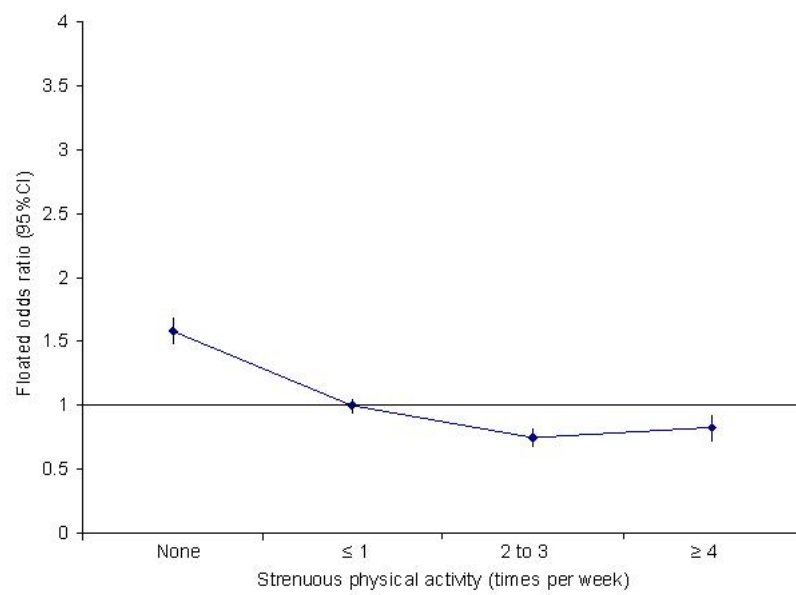


FIGURE 5.6: Odds ratio of diabetes by category of strenuous physical activity in a model adjusted for age, region, and socioeconomic group (OR, 95%CI)

TABLE 5.4: Odds ratio of diabetes among postmenopausal women associated with parity and breastfeeding in models adjusted for various important co-factors (OR, 95%CI)

		Unadjusted Results			Model A - adjusted for age, region, socioeconomic group, parity, and breastfeeding as appropriate				Model A + additionally adjusted for BMI		Model A + additionally adjusted for smoking		Model A + additionally adjusted for BMI and smoking		Model A + additionally adjusted BMI, smoking, alcohol, and strenuous physical activity	
		Mean	Cases	Risk (%)	OR	95%CI	OR	95%CI	OR	95%CI	OR	95%CI	OR	95%CI	OR	95%CI
All women																
Parity																
Nulliparous women		0	909	1.42	1.00	-	1.00	-	1.00	-	1.00	-	1.00	-	1.00	-
Parous women		2	7,030	1.53	1.07	(0.87 - 1.33)	1.22	(1.10 - 1.33)	1.09	(1.00 - 1.19)	1.21	(1.11 - 1.32)	1.08	(0.99 - 1.18)	1.05	(0.96 - 1.14)
Ever breastfed*																
No		0	2,257	1.73	1.00	(0.82 - 1.22)	1.00	(0.98 - 1.02)	1.00	(0.98 - 1.02)	1.00	(0.98 - 1.02)	1.00	(0.98 - 1.02)	1.00	(0.98 - 1.02)
Yes		4.9	4,773	1.45	0.87	(0.71 - 1.06)	0.78	(0.74 - 0.82)	0.85	(0.80 - 0.90)	0.78	(0.74 - 0.83)	0.86	(0.81 - 0.91)	0.89	(0.84 - 0.94)
Parity*																
1		1	1,019	1.45	1.17	(1.10 - 1.25)	1.07	(0.94 - 1.22)	1.08	(1.01 - 1.17)	1.06	(0.99 - 1.14)	1.07	(1.00 - 1.16)	1.07	(0.99 - 1.15)
2		2	2,833	1.24	1.00	(0.96 - 1.04)	1.00	(0.89 - 1.12)	1.00	(0.97 - 1.03)	1.00	(0.98 - 1.02)	1.00	(0.97 - 1.03)	1.00	(0.96 - 1.04)
3		3	1,862	1.64	1.33	(1.27 - 1.39)	1.30	(1.14 - 1.47)	1.18	(1.11 - 1.26)	1.29	(1.22 - 1.37)	1.17	(1.10 - 1.25)	1.16	(1.09 - 1.23)
>4		4.4	1,316	2.50	2.05	(1.94 - 2.17)	1.85	(1.62 - 2.11)	1.45	(1.35 - 1.56)	1.82	(1.70 - 1.96)	1.42	(1.32 - 1.53)	1.37	(1.28 - 1.47)
Test for trend					p<0.0001		p<0.0001		p<0.0001		p<0.0001		p<0.0001		p<0.0001	
Total duration of breastfeeding*																
Did not breastfeed		0	2,257	1.70	1.00	(0.96 - 1.04)	1.00	(0.89 - 1.12)	1.00	(0.97 - 1.03)	1.00	(0.98 - 1.02)	1.00	(0.97 - 1.03)	1.00	(0.96 - 1.04)
< 6 months		2.3	2,556	1.47	0.86	(0.83 - 0.90)	0.82	(0.72 - 0.93)	0.87	(0.83 - 0.92)	0.82	(0.78 - 0.87)	0.88	(0.83 - 0.93)	0.90	(0.85 - 0.95)
6 - 9 months		7.2	858	1.39	0.81	(0.76 - 0.87)	0.77	(0.67 - 0.88)	0.84	(0.77 - 0.91)	0.77	(0.71 - 0.83)	0.85	(0.79 - 0.92)	0.88	(0.81 - 0.96)
10 + months		18.6	1,359	1.38	0.81	(0.77 - 0.85)	0.70	(0.61 - 0.80)	0.80	(0.75 - 0.86)	0.71	(0.66 - 0.76)	0.84	(0.76 - 0.88)	0.87	(0.81 - 0.93)
Test for trend				p = 0.09			p=0.01		p=0.07		p=0.07		p=0.04		p=0.06	
* Parous women only																

* Parous women only

TABLE 5.5: χ^2 statistic for heterogeneity across categories of parity (three degrees on freedom) and durations of breastfeeding (three degrees of freedom) for each of the models in table 5.4 (page 152)

Chi-square for heterogeneity across categories		
	Parity [3 degrees freedom]	Duration breastfeeding [3 degrees freedom]
Model A - Adjusted for age, socioeconomic group, parity, and breastfeeding as appropriate, and stratified by region	302.82	104.97
Model A + additionally adjusted for BMI	108.63	41.73
Model A + additionally adjusted for smoking	290.23	97.1
Model A + additionally adjusted for BMI and smoking	97.32	35.28
Model A + additionally adjusted BMI, smoking, alcohol, and strenuous physical activity	83.83	25.44

TABLE 5.6: Odds ratio of diabetes among parous postmenopausal women associated with parity and breastfeeding in a model adjusted for age, region, smoking, socioeconomic group, BMI, alcohol, and strenuous physical activity (OR, 95% CI)

Characteristics of women	Cases	Population at risk	Risk (%)	OR	(95% CI)
Parity					
0	909	65,090	1.40	1.03	(0.96 - 1.11)
1	1,019	70,376	1.45	1.07	(0.99 - 1.15)
2	2,833	229,285	1.24	1.00	(0.96 - 1.04)
3	1,862	113,845	1.64	1.16	(1.09 - 1.23)
≥ 4	1,316	52,605	2.50	1.36	(1.28 - 1.47)
Breastfeeding duration					
0	2,257	132,611	1.70	1.00	(0.96 - 1.04)
< 6	2,556	173,534	1.47	0.90	(0.85 - 0.95)
6 to 10	858	61,638	1.39	0.88	(0.81 - 0.96)
≥ 10	1,359	98,328	1.38	0.87	(0.81 - 0.93)
Socioeconomic group					
Upper	1,318	116,201	1.13	0.95	(0.90 - 1.01)
	1,386	114,387	1.21	0.94	(0.89 - 0.99)
Middle	1,512	109,650	1.38	1.00	(0.95 - 1.05)
	1,740	101,474	1.71	1.10	(1.05 - 1.16)
Lowest	1,916	85,287	2.25	1.17	(1.11 - 1.23)
Body mass index					
<20	58	18,947	0.31	0.59	(0.45 - 0.76)
20 to 22.4	208	83,046	0.25	0.51	(0.44 - 0.60)
22.5 to 24.9	714	144,049	0.50	1.00	(0.95 - 1.05)
25 to 27.4	1,169	113,091	1.03	1.99	(1.85 - 2.15)
27.5 to 29.9	1,400	70,405	1.99	3.72	(3.46 - 4.01)
30 to 32.4	1,238	37,749	3.28	6.08	(5.63 - 6.57)
32.5 to 34.9	1,005	21,124	4.76	8.71	(8.02 - 9.46)
≥ 35	1,663	21,429	7.76	14.39	(8.02 - 9.46)
Smoking					
Never	3,634	278,191	1.31	1.00	(0.96 - 1.04)
Current <5	97	7,808	1.24	1.15	(0.94 - 1.41)
Current 5 to 9	209	13,380	1.56	1.49	(1.36 - 1.62)
Current 10 to 14	339	19,793	1.71	1.54	(1.37 - 1.72)
Current 15 to 19	299	18,304	1.63	1.38	(1.22 - 1.56)
Current 20 to 24	287	14,028	2.05	1.53	(1.35 - 1.73)
Current - ≥ 25	111	4,118	2.70	1.87	(1.53 - 2.27)
Strenuous physical activity					
None	4,616	230,559	2.00	1.25	(1.22 - 1.29)
1 or less per week	1,999	170,730	1.17	1.00	(0.96 - 1.04)
2 to 3 times	744	84,096	0.88	0.84	(0.78 - 0.90)
4 or more times	311	31,625	0.98	0.90	(0.80 - 1.01)

TABLE 5.7: Odds ratio of diabetes among parous postmenopausal women associated with increasing parity and duration of breastfeeding stratified by category of BMI (kg/m^2), (OR, 95% CI)

Body mass index strata		< 25 kg/m^2		25 to 30 kg/m^2		30 to 35 kg/m^2		> 35 kg/m^2	
		Risk (%)	OR (95%CI)	Risk (%)	OR (95%CI)	Risk (%)	OR (95%CI)	Risk (%)	OR (95%CI)
Parity*	1	0.44	1.20 (1.02 - 1.41)	1.38	1.05 (0.93 - 1.17)	3.70	1.05 (0.96 - 1.14)	7.97	1.14 (0.98 - 1.32)
	2	0.33	1.00 (0.90 - 1.11)	1.20	1.00 (0.94 - 1.07)	3.57	1.00 (0.95 - 1.05)	6.86	1.00 (0.91 - 1.09)
	3	0.43	1.24 (1.08 - 1.42)	1.48	1.19 (1.09 - 1.29)	3.91	1.16 (1.09 - 1.23)	8.69	1.24 (1.12 - 1.38)
	>4	0.73	1.86 (1.56 - 2.22)	2.12	1.53 (1.37 - 1.70)	4.92	1.30 (1.20 - 1.40)	9.12	1.17 (1.03 - 1.34)
Total duration of breastfeeding*									
Did not breastfeed		0.47	1.00 (0.88 - 1.14)	1.51	1.00 (0.92 - 1.08)	3.96	1.00 (0.94 - 1.06)	7.86	1.00 (0.90 - 1.11)
< 6 months		0.38	0.76 (0.67 - 0.85)	1.35	0.86 (0.80 - 0.92)	4.02	0.91 (0.91 - 1.01)	7.70	1.00 (0.91 - 1.09)
6 - 9 months		0.40	0.82 (0.68 - 0.99)	1.32	0.83 (0.74 - 0.93)	3.75	0.85 (0.85 - 1.02)	7.80	1.04 (0.89 - 1.21)
10 + months		0.38	0.76 (0.64 - 0.88)	1.42	0.86 (0.78 - 0.95)	3.58	0.81 (0.81 - 0.95)	8.17	1.07 (0.94 - 1.23)

* Parous women only

** Adjusted for age, region, socioeconomic group, smoking, alcohol, strenuous physical activity, parity, and breastfeeding as appropriate

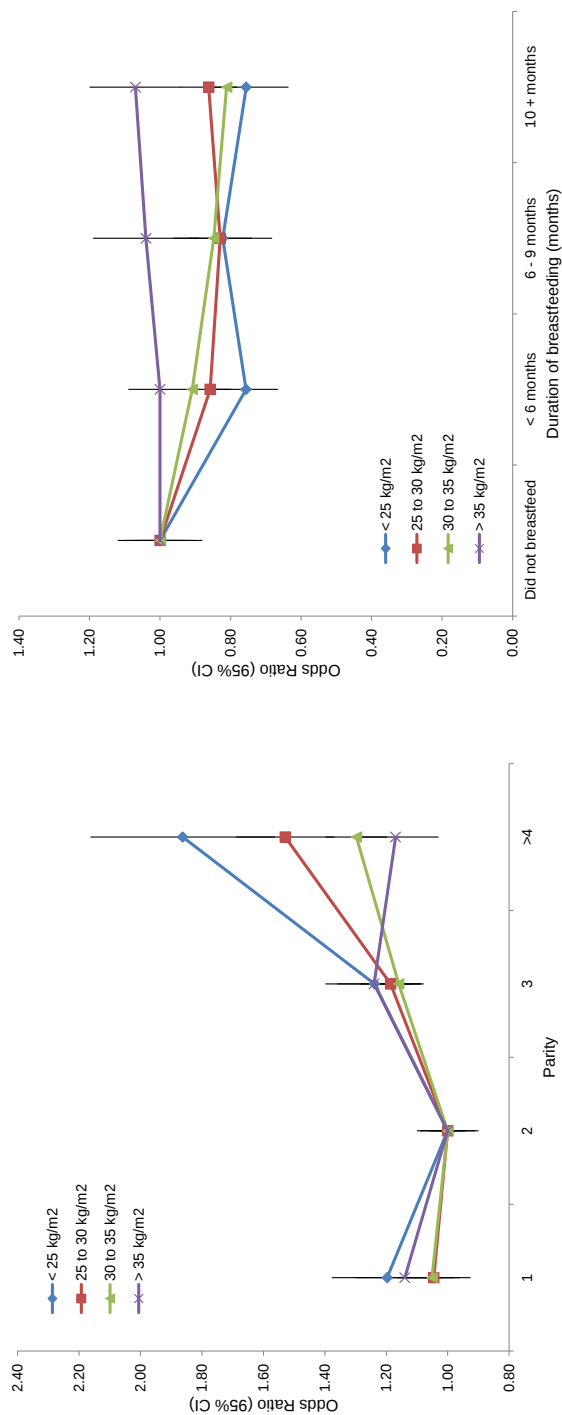


FIGURE 5.7: Interaction effects of BMI (kg/m^2) on the odds ratios of diabetes associated with increasing parity and duration of breastfeeding (OR, 95% CI)

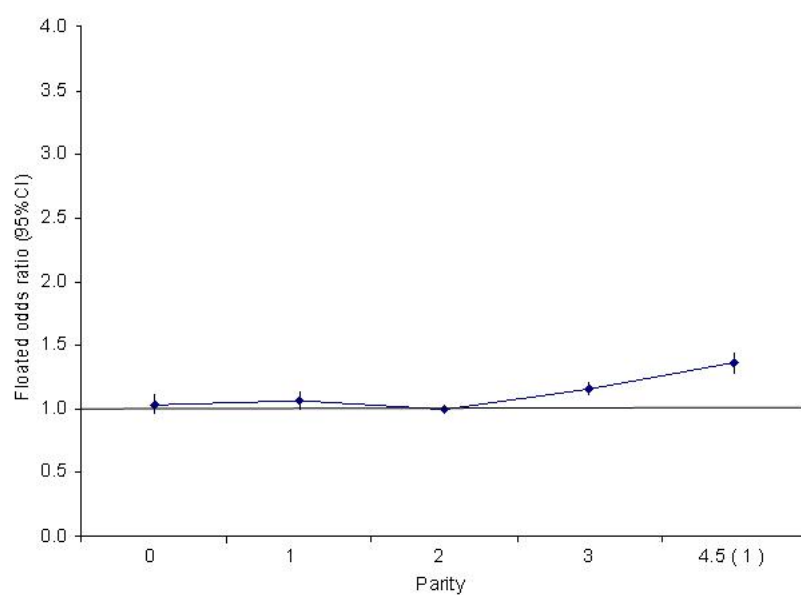


FIGURE 5.8: Odds ratio of diabetes by parity in the fully adjusted model (i.e. adjusted for age, region, smoking, BMI, alcohol, strenuous physical activity, and duration of breastfeeding as appropriate), (OR, 95% CI)

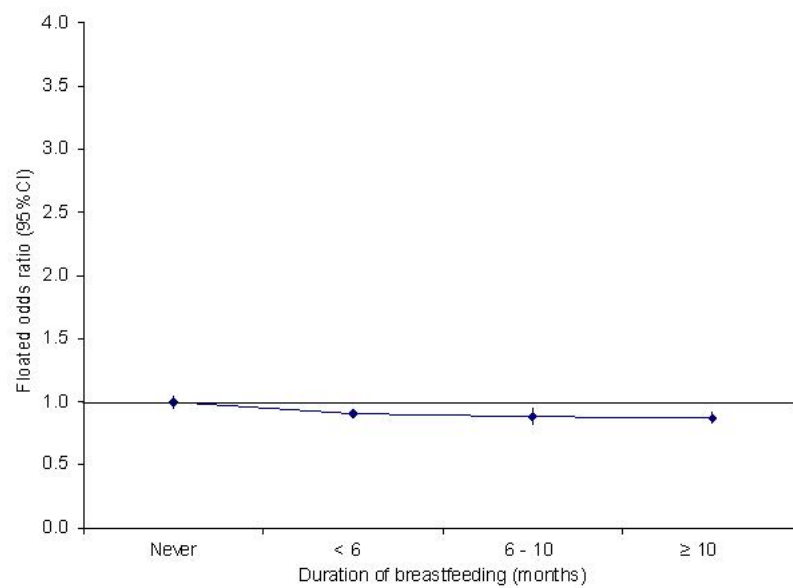


FIGURE 5.9: Odds ratio of diabetes among parous postmenopausal women by breastfeeding in the fully adjusted model, (OR, 95% CI)

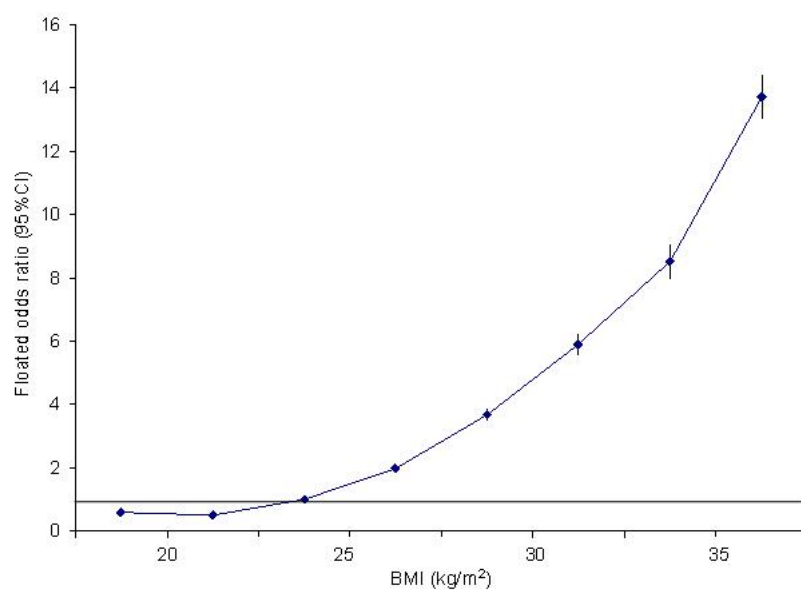


FIGURE 5.10: Odds ratio of diabetes by BMI (kg/m^2) in the fully adjusted model, (OR, 95% CI)

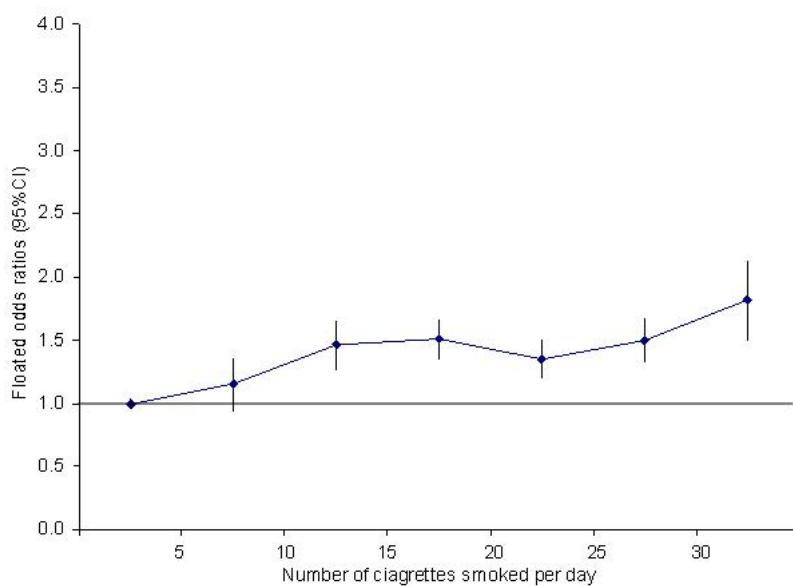


FIGURE 5.11: Odds ratio of diabetes by smoking in the fully adjusted model, (OR, 95% CI)

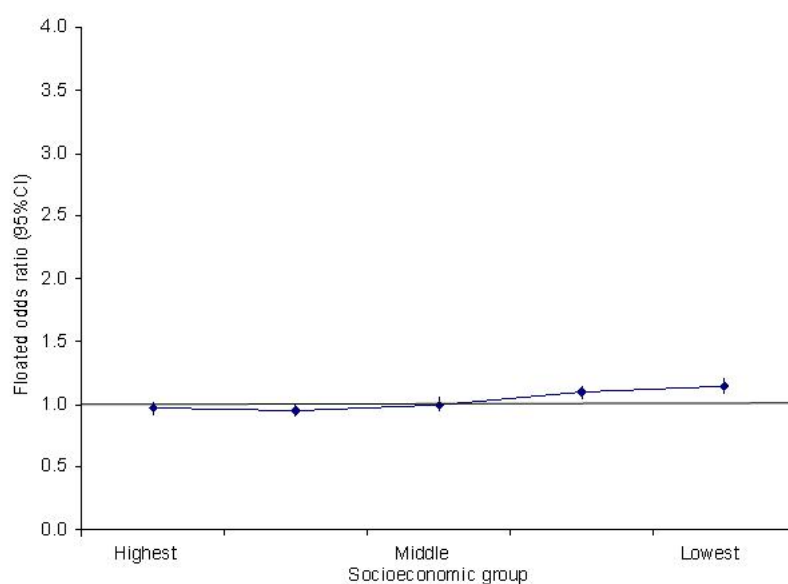


FIGURE 5.12: Odds ratio of diabetes by socioeconomic group in the fully adjusted model, (OR, 95% CI)

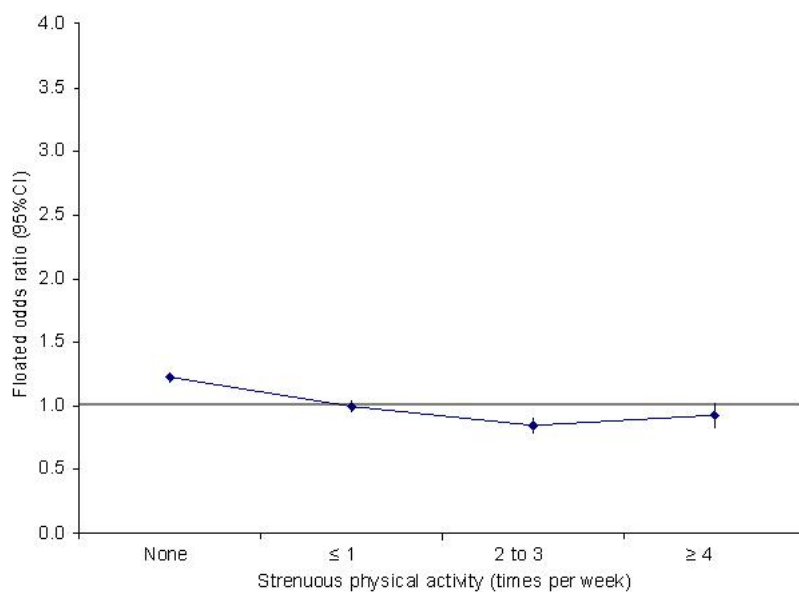


FIGURE 5.13: Odds ratio of diabetes by strenuous physical activity in the fully adjusted model, (OR, 95% CI)

TABLE 5.8: Odds ratio of diabetes among postmenopausal women associated with parity and duration of breastfeeding in the fully adjusted model and in models which have additional adjustments made for ethnicity, family history of diabetes, and mother's own birth-weight (kg), (OR, 95% CI)

	Mean	Model B - Final model**	Model B + additional adjustment for ethnicity	Model B + additional adjustment for a family history of diabetes	Model B + additional adjustment for mother's own birth weight
		OR (95%CI)	OR (95%CI)	OR (95%CI)	OR (95%CI)
Parity					
Nulliparous women	0	1.00	1.00	1.00	1.00
Parous women	2	1.06 (0.96 - 1.14)	1.06 (0.97 - 1.15)	1.05 (0.97 - 1.15)	1.04 (0.96 - 1.14)
Ever breastfed*					
No	0	1.00 (0.98 - 1.02)	1.00 (0.98 - 1.02)	1.00 (0.98 - 1.02)	1.00 (0.98 - 1.02)
Yes	4.9	0.89 (0.84 - 0.94)	0.87 (0.83 - 0.93)	0.88 (0.83 - 0.94)	0.89 (0.94 - 0.94)
Parity*					
1	1	1.07 (0.99 - 1.15)	1.06 (0.99 - 1.13)	1.06 (0.99 - 1.13)	1.06 (1.00 - 1.13)
2	2	1.00 (0.97 - 1.03)	1.00 (0.97 - 1.03)	1.00 (0.97 - 1.03)	1.00 (0.97 - 1.03)
3	3	1.16 (1.09 - 1.24)	1.16 (1.11 - 1.22)	1.16 (1.11 - 1.22)	1.16 (1.10 - 1.22)
>4	4.4	1.37 (1.27 - 1.48)	1.35 (1.27 - 1.44)	1.37 (1.29 - 1.46)	1.36 (1.28 - 1.45)
Total duration of breastfeeding*					
Did not breastfeed	0	1.00 (0.97 - 1.03)	1.00 (0.95 - 1.05)	1.00 (0.95 - 1.05)	1.00 (0.95 - 1.05)
< 6 months	2.3	0.90 (0.84 - 0.95)	0.89 (0.84 - 0.95)	0.89 (0.84 - 0.95)	0.90 (0.86 - 0.94)
6 - 9 months	7.2	0.88 (0.81 - 0.96)	0.87 (0.81 - 0.93)	0.88 (0.82 - 0.95)	0.89 (0.83 - 0.95)
10 + months	18.6	0.87 (0.80 - 0.93)	0.83 (0.78 - 0.89)	0.86 (0.81 - 0.92)	0.87 (0.83 - 0.93)

* Parous women only

** Adjusted for age, socioeconomic group, BMI, smoking, alcohol, strenuous physical activity, region, parity, and breastfeeding as appropriate

TABLE 5.9: Odds ratio of diabetes among parous postmenopausal women associated with parity and duration of breastfeeding in the fully adjusted model (age, region, socioeconomic group, BMI, smoking, alcohol, strenuous physical activity, and parity and breastfeeding as appropriate) and in a model which is similarly adjusted but where the category mean for BMI was calculated using measured heights and weights from a random sample of women (taken about 8 years after recruitment), (OR, 95% CI)

Source of BMI data	Self-report at recruitment		From measured heights and weights***	
	OR	(95%CI)	OR	(95%CI)
Parity*				
0	1.03	(0.96 - 1.11)	1.03	(0.96 - 1.11)
1	1.07	(1.00 - 1.13)	1.07	(1.00 - 1.13)
2	1.00	(0.96 - 1.04)	1.00	(0.96 - 1.04)
3	1.16	(1.11 - 1.22)	1.16	(1.11 - 1.22)
>4	1.36	(1.28 - 1.45)	1.36	(1.28 - 1.45)
Total duration of breastfeeding*				
Did not breastfeed	1.00	(0.96 - 1.04)	1.00	(0.96 - 1.04)
< 6 months	0.90	(0.86 - 0.94)	0.90	(0.86 - 0.94)
6 - 9 months	0.89	(0.83 - 0.95)	0.89	(0.83 - 0.95)
10 + months	0.87	(0.82 - 0.92)	0.87	(0.82 - 0.92)

* Parous women only

** Adjusted for age, socioeconomic group, smoking, alcohol, strenuous physical activity, region, parity, and breastfeeding as appropriate

*** Random sample of 3,745 women with heights and weights measured about 8 years after recruitment

Bibliography

- [1] King H, Aubert RE, Herman WH. Global burden of diabetes, 1995-2025: prevalence, numerical estimates, and projections. *Diabetes Care*. 1998 Sep;21(9):1414–1431.
- [2] Zimmet P, Alberti KGMM, Shaw J. Global and societal implications of the diabetes epidemic. *Nature*. 2001 Dec;414(6865):782–787.
- [3] Bagust A, Hopkinson PK, Maslove L, Currie CJ. The projected health care burden of Type 2 diabetes in the UK from 2000 to 2060. *Diabetic Medicine*. 2002;19(s4):1–5.
- [4] González ELM, Johansson S, Wallander M, Rodríguez LAG. Trends in the prevalence and incidence of diabetes in the UK: 1996–2005. *Journal of Epidemiology and Community Health*. 2009 Apr;63(4):332–336.
- [5] Department of Health. NSF Diabetes: What is diabetes? Crown Copyright; 2010. This document, the first part of the National Service Framework for Diabetes, sets out twelve new standards and the key interventions necessary to raise the standards of diabetes care.
- [6] Colditz GA, Willett WC, Rotnitzky A, Manson JE. Weight Gain as a Risk Factor for Clinical Diabetes Mellitus in Women. *Annals of Internal Medicine*. 1995 Apr;122(7):481–486.
- [7] Pontiroli AE, Galli L. Duration of obesity is a risk factor for non-insulin-dependent diabetes mellitus, not for arterial hypertension or for hyperlipidaemia. *Acta Diabetologica*. 1998 Oct;35(3):130–136.
- [8] Connolly V, Unwin N, Sherriff P, Bilous R, Kelly W. Diabetes prevalence and socio-economic status: a population based study showing increased prevalence of type 2 diabetes mellitus in deprived areas. *Journal of Epidemiology and Community Health*. 2000 Mar;54(3):173–177.
- [9] Evans JMM, Newton RW, Ruta DA, MacDonald TM, Morris AD. Socio-economic status, obesity and prevalence of Type 1 and Type 2 diabetes mellitus. *Diabetic Medicine*. 2000;17(6):478–480.
- [10] Weinstein AR, Sesso HD, Lee IM, Cook NR, Manson JE, Buring JE, et al. Relationship of Physical Activity vs Body Mass Index With Type 2 Diabetes in Women. *JAMA: The Journal of the American Medical Association*. 2004;292(10):1188–1194.
- [11] Kumari M, Head J, Marmot M. Prospective Study of Social and Other Risk Factors for Incidence of Type 2 Diabetes in the Whitehall II Study. *Arch Intern Med*. 2004 Sep;164(17):1873–1880.

- [12] Spencer EA, Pirie KL, Stevens RJ, Beral V, Brown A, Liu B, et al. Diabetes and modifiable risk factors for cardiovascular disease: the prospective Million Women Study. *European Journal of Epidemiology*. 2008;23(12):793–799.
- [13] Hu FB, Manson JE, Stampfer MJ, Colditz G, Liu S, Solomon CG, et al. Diet, Lifestyle, and the Risk of Type 2 Diabetes Mellitus in Women. *New England Journal of Medicine*. 2001;345(11):790–797.
- [14] Chaturvedi N. The burden of diabetes and its complications: Trends and implications for intervention. *Diabetes Research and Clinical Practice*. 2007 May;76(3, Supplement 1):S3–S12.
- [15] World Health Organisation W. Global Database on Body Mass Index [online]; 2010. Internet [cited 02 November 2010]. Available from: http://apps.who.int/bmi/index.jsp?introPage=intro_3.html.
- [16] Beral V. Long term effects of childbearing on health. *J Epidemiol Community Health*. 1985 Dec;39(4):343–346.
- [17] Manson JE, Rimm EB, Colditz GA, Stampfer MJ, Willett WC, Arky RA, et al. Parity and incidence of non-insulin-dependent diabetes mellitus. *The American Journal of Medicine*. 1992 Jul;93(1):13–18.
- [18] Hanley AJG, McKeown-Eyssen G, Harris SB, Hegele RA, Wolever TMS, Kwan J, et al. Association of parity with risk of type 2 diabetes and related metabolic disorders. *Diabetes Care*. 2002 Apr;25(4):690–695.
- [19] Nicholson WK, Asao K, Brancati F, Coresh J, Pankow JS, Powe NR. Parity and risk of type 2 diabetes: the Atherosclerosis Risk in Communities Study. *Diabetes Care*. 2006 Nov;29(11):2349–2354.
- [20] Gunderson EP, Lewis CE, Tsai A, Chiang V, Carnethon M, Quesenberry CP, et al. A 20-Year Prospective Study of Childbearing and Incidence of Diabetes in Young Women, Controlling for Glycemia Before Conception: The Coronary Artery Risk Development in Young Adults (CARDIA) Study. *Diabetes*. 2007 Dec;56(12):2990–2996.
- [21] Liu B, Jorm L, Banks E. Parity, breastfeeding, and the subsequent risk of maternal type 2 diabetes. *Diabetes Care*. 2010 Jun;33(6):1239–1241.
- [22] Stuebe AM, Rich-Edwards JW, Willett WC, Manson JE, Michels KB. Duration of Lactation and Incidence of Type 2 Diabetes. *JAMA*. 2005 Nov;294(20):2601–2610.
- [23] Villegas R, Gao Y, Yang G, Li HL, Elasy T, Zheng W, et al. Duration of breast-feeding and the incidence of type 2 diabetes mellitus in the Shanghai Women's Health Study. *Diabetologia*. 2008 Feb;51(2):258–266.
- [24] Schwarz EB, Brown JS, Creasman JM, Stuebe A, McClure CK, Eeden SKVD, et al. Lactation and Maternal Risk of Type 2 Diabetes: A Population-based Study. *The American Journal of Medicine*. 2010;123(9):863.e1–863.e6.
- [25] The Million Women Study Collaborative Group. The Million Women Study: design and characteristics of the study population. *Breast Cancer Research: BCR*. 1999;1(1):73–80.

- [26] Banks E, Beral V, Cameron R, Hogg A, Langley N, Barnes I, et al. Agreement between general practice prescription data and self-reported use of hormone replacement therapy and treatment for various illnesses. *Journal of Epidemiology and Biostatistics*. 2001;6(4):357.
- [27] Heliövaara M, Aromaa A, Klaukka T, Knekt P, Joukamaa M, Impivaara O. Reliability and validity of interview data on chronic diseases The mini-Finland health survey. *Journal of Clinical Epidemiology*. 1993 Feb;46(2):181–191.
- [28] Norrish A, North D, Kirkman P, Jackson R. Validity of Self-reported Hospital Admission in a Prospective Study. *American Journal of Epidemiology*. 1994 Nov;140(10):938–942.
- [29] Bowlin SJ, Morrill BD, Nafziger AN, Lewis C, Pearson TA. Reliability and changes in validity of self-reported cardiovascular disease risk factors using dual response: The behavioral risk factor survey. *Journal of Clinical Epidemiology*. 1996 May;49(5):511–517.
- [30] The Italian Longitudinal Study on Aging Working Group. Prevalence of chronic diseases in older Italians: comparing self-reported and clinical diagnoses. *International Journal of Epidemiology*. 1997 Oct;26(5):995–1002.
- [31] Goldman N, Lin I, Weinstein M, Lin Y. Evaluating the quality of self-reports of hypertension and diabetes. *Journal of Clinical Epidemiology*. 2003 Feb;56(2):148–154.
- [32] Okura Y, Urban LH, Mahoney DW, Jacobsen SJ, Rodeheffer RJ. Agreement between self-report questionnaires and medical record data was substantial for diabetes, hypertension, myocardial infarction and stroke but not for heart failure. *Journal of Clinical Epidemiology*. 2004 Oct;57(10):1096–1103.
- [33] Simpson CF, Boyd CM, Carlson MC, Griswold ME, Guralnik JM, Fried LP. Agreement Between Self-Report of Disease Diagnoses and Medical Record Validation in Disabled Older Women: Factors That Modify Agreement. *Journal of the American Geriatrics Society*. 2004;52(1):123–127.
- [34] Townsend P, Phillimore P, Beattie A. *Health and deprivation*. Routledge; 1988.
- [35] Alberti KG, Zimmet PZ. Definition, diagnosis and classification of diabetes mellitus and its complications. Part 1: diagnosis and classification of diabetes mellitus provisional report of a WHO consultation. *Diabetic Medicine: A Journal of the British Diabetic Association*. 1998 Jul;15(7):539–553.
- [36] Hu F. *Obesity epidemiology*. Oxford: Oxford University Press US; 2008.
- [37] MacMahon S, Peto R, Collins R, Godwin J, MacMahon S, Cutler J, et al. Blood pressure, stroke, and coronary heart disease : Part 1, prolonged differences in blood pressure: prospective observational studies corrected for the regression dilution bias. *The Lancet*. 1990 Mar;335(8692):765–774.
- [38] Easton DF, Peto J, Babiker AGAG. Floating absolute risk: An alternative to relative risk in survival and case-control analysis avoiding an arbitrary reference group. *Statistics in Medicine*. 1991 Jul;10(7):1025–1035.
- [39] Biggs ML, Mukamal KJ, Luchsinger JA, Ix JH, Carnethon MR, Newman AB, et al. Association Between Adiposity in Midlife and Older Age and Risk of Diabetes in Older Adults. *JAMA*. 2010 Jun;303(24):2504–2512.

- [40] Fletcher C, Peto R, Speizer F, Tinker C. A follow-up study of the natural history of obstructive bronchitis. Oxford University Press; 1976.
- [41] Spencer EA, Appleby PN, Davey GK, Key TJ. Validity of Self-Reported Height and Weight in 4808 EPIC Oxford Participants. *Public Health Nutrition*. 2002;5(04):561–565.
- [42] Boyko EJ, Alderman BW, Keane EM, Baron AE. Effects of childbearing on glucose tolerance and NIDDM prevalence. *Diabetes Care*. 1990;13(8):848–854.
- [43] Collins VR, Dowse GK, Zimmet PZ. Evidence against association between parity and NIDDM from five population groups. *Diabetes Care*. 1991 Nov;14(11):975–981.
- [44] Manson JE, Willett WC, Stampfer MJ, Colditz GA, Hunter DJ, Hankinson SE, et al. Body Weight and Mortality among Women. *N Engl J Med*. 1995 Sep;333(11):677–685.
- [45] Carey VJ, Walters EE, Colditz GA, Solomon CG, Willett WC, Rosner BA, et al. Body Fat Distribution and Risk of Non-Insulin-dependent Diabetes Mellitus in Women. *American Journal of Epidemiology*. 1997 Apr;145(7):614–619.
- [46] Meisinger C, Doring A, Thorand B, Heier M, Lowel H. Body fat distribution and risk of type 2 diabetes in the general population: are there differences between men and women? The MONICA/KORA Augsburg Cohort Study. *Am J Clin Nutr*. 2006 Sep;84(3):483–489.
- [47] Colditz G, Willett WC, Stamler MJ, Manson JE, Hennekens CH, Arky RA, et al. Weight as a risk factor for clinical diabetes in women. *American Journal of Epidemiology*. 1990;132(3):501–513.
- [48] Hart CL, Hole DJ, Lawlor DA, Smith GD. How many cases of Type 2 diabetes mellitus are due to being overweight in middle age? Evidence from the Midspan prospective cohort studies using mention of diabetes mellitus on hospital discharge or death records. *Diabetic Medicine: A Journal of the British Diabetic Association*. 2007 Jan;24(1):73–80.
- [49] Taylor AE, Ebrahim S, Ben-Shlomo Y, Martin RM, Whincup PH, Yarnell JW, et al. Comparison of the associations of body mass index and measures of central adiposity and fat mass with coronary heart disease, diabetes, and all-cause mortality: a study using data from 4 UK cohorts. *The American Journal of Clinical Nutrition*. 2010;.
- [50] Perry IJ, Wannamethee SG, Walker MK, Thomson AG, Whincup PH, Shaper AG. Prospective study of risk factors for development of non-insulin dependent diabetes in middle aged British men. *BMJ*. 1995 Mar;310(6979):560–564.
- [51] Rimm EB, Chan J, Stampfer MJ, Colditz GA, Willett WC. Prospective study of cigarette smoking, alcohol use, and the risk of diabetes in men. *BMJ*. 1995 Mar;310(6979):555–559.
- [52] Hu FB, Sigal RJ, Rich-Edwards JW, Colditz GA, Solomon CG, Willett WC, et al. Walking Compared With Vigorous Physical Activity and Risk of Type 2 Diabetes in Women: A Prospective Study. *JAMA*. 1999 Oct;282(15):1433–1439.
- [53] Fuller JH, Stevens LK, Wang S. Risk factors for cardiovascular mortality and morbidity: The WHO multinational study of vascular disease in diabetes. *Diabetologia*. 2001;44(14):S54–S64.

- [54] Promislow JHE, Gladen BC, Sandler DP. Maternal Recall of Breastfeeding Duration by Elderly Women. *Am J Epidemiol*. 2005 Feb;161(3):289–296.
- [55] Ko K, Park S, Kim Y, Bae J, Jun J, Gwack J, et al. Reliability of a questionnaire for women's reproductive history. *Journal of Preventive Medicine and Public Health Korean*. 2008;41(3).

Chapter 6

The effects of childbearing on women's risk of ischaemic heart disease after the menopause

Abstract

Objectives: *To explore the association between childbearing patterns and women's risk of ischaemic heart disease after the menopause*

Design: *Prospective cohort analysis of the risk of incident ischaemic heart disease associated with women's childbearing histories*

Setting: *Population-based cohort of UK women recruited in 1996 to 2001 through the National Health Service Breast Screening Programme (the Million Women Study)*

Participants: *869,828 women who at recruitment were postmenopausal and who reported their reproductive and medical histories, as well as other personal and socio-demographic factors*

Main outcome measure: *Risk of ischaemic heart disease (IHD) and hazard ratios (HR) adjusted for age, socioeconomic status, body mass index, smoking, alcohol, and physical activity, and stratified by geographic region*

Results: *Women were aged 57.5 years [SD 4] on average and had a mean BMI of 26.2 kg/m² [SD 5]. Women had an IHD incidence rate of 4.6 per 1000 person years (CI 4.5 to 4.7). 88% of the study population were parous, with mean parity of 2.1 children [SD 1.2]. Of parous women 60% had breastfed and the average total duration of breastfeeding was 4.7 months [SD 7.9]. In the unadjusted analyses, the effects of parity and breastfeeding on women's risk of IHD were similar in size to the effects of socioeconomic status, BMI and strenuous physical activity. Of the selected co-factors smoking was most strongly associated with the risk of IHD. After adjusting for age, socioeconomic group, BMI, smoking, alcohol, strenuous physical activity, and geographic region nulliparous women had the lowest risk of incident IHD (HR 0.90, CI 0.87 to 0.93). Among parous women increasing parity was associated with an increase in the risk of incident IHD; As compared to women of parity ≤ 2 women with parity three had a 12% increase in risk (HR 1.12, CI 1.09 to 1.16) while women of parity four or more had a 31% increase in risk (HR 1.31, CI 1.27 to 1.36)). After adjustment parous women with a history of ever breastfeeding had an 8% lower risk of incident IHD (HR 0.92, CI 0.90 to 0.95) and a small decreasing trend was seen with longer durations of breastfeeding (HR 0.96 per month of breastfeeding, CI 0.95 to 0.98). Smoking and body mass index were both strongly associated with an increased risk in incident IHD and in contrast, the effects of parity and breastfeeding on women's risk of IHD were small.*

Conclusions: *In this analysis of over 35,000 events in over 800,000 women I found parity of three or more was associated with an increase in a woman's risk of IHD after the menopause. Among parous women having ever breastfed offset this risk slightly, however the effects of childbearing were small in comparison to known important risk factors. For example smoking less or stopping completely and having a BMI less than 25 kg/m², were more important than either parity or breastfeeding history in altering women's risk of ischaemic heart disease after the menopause.*

6.1 Introduction

Ischaemic heart disease is a major cause of premature death among postmenopausal women in developed countries. [1] A large amount of data exists on risk factors for heart disease and there is little evidence to suggest differences between men and women for factors they share. [2–6]

It has been suggested that factors unique to women (such as parity and breastfeeding) may also influence risk. Results from several small cross-sectional analyses suggest that increasing parity is associated with an increased likelihood of ischaemic heart disease and that for parous women having breastfed may be protective against ischaemic heart disease. [7–11] Few prospective studies exist that are able to look at the long-term effects of childbearing and breastfeeding on women's risk of heart disease after the menopause.

Investigators from a large population-based cohort of Swedish women found a modest association between increasing parity and incident heart disease but they did not have information on breastfeeding and were not able to adjust for smoking. [12] Two relatively small prospective studies from the USA investigating the relation between women's breastfeeding histories and their subsequent risk of incident heart disease have reported conflicting results. [13, 14] Findings reported from the Women's Health Initiative study, which included both a cross-sectional and prospective analysis, suggest that while women who had breastfed were less likely to have prevalent heart disease at study entry there was no effect of having breastfed on women's risk of incident heart disease during 9 years of follow-up. [13] In contrast, investigators from the Nurses Health Study reported that in parous women longer durations of breastfeeding were associated with a lower risk of incident heart disease (myocardial infarction)

over 15 years of follow-up. [14] To my knowledge no study has looked at the independent (and joint) effects of parity and breastfeeding on women's risk of ischaemic heart disease, and compared the effect size to that of well established risk factors.

In this chapter I report on the risk of incident ischaemic heart disease within a large cohort of postmenopausal UK women in relation to their childbearing histories accounting for age, region, socioeconomic group, body mass index, smoking, alcohol, and strenuous physical activity. I also report on size of the effects of childbearing in relation to smoking, BMI, socioeconomic group, and strenuous physical activity.

6.2 Methods

6.2.1 Study Population

Recruitment

The Million Women Study is a prospective study of women living in England and Scotland. Between 1996 and 2001, 1.3 million women aged 50 to 64 years were recruited into the cohort through the National Health System (NHS) Breast Screening Programme. At recruitment women completed an initial questionnaire which asked about reproductive and medical histories, socioeconomic and lifestyle factors, and other personal characteristics.

Repeat survey

Starting in 1999 the first round of repeat surveys were mailed to study participants about 3 years after they were recruited (65% response rate). At repeat survey women

were asked about new exposures and information obtained at recruitment on lifestyle and other characteristics including smoking and alcohol exposure was updated.

Follow-up

Participants are routinely followed up through linkage to centrally held computerised records using their unique health care number (NHS number), date of birth and other identifying details. The linked records include NHS central registries (which contain detailed records for deaths, cancers, and emigrations), [15, 16] the Scottish morbidity records in Scotland (hospital admissions data), and the Hospital Episode Statistics in England (hospital admissions data). [17, 18] The hospital admissions databases contain a record of all NHS in-patient admissions from January 1981 in Scotland and April 1997 in England. [18, 19] Diagnoses at hospital discharge (including a main diagnosis and up to 13 additional diagnoses) or for deaths (the underlying cause of death) are coded using ICD-10 (the International Statistical Classification of Diseases and Related Health Problems 10th Revision, ICD-10). [20]

Full details of the study design and methods have been published elsewhere, [21] aspects of the study design and methods relevant to this thesis are described in chapter 2 (page 18). Study questionnaires can be viewed at <http://www.millionwomen.org> and have been included in this thesis in Appendix A, page 241. The study was approved by the Anglia and Oxford Multi-Centre Research Ethics Committee. Access and linkage to hospital records was approved by the Information Centre for Health and Social Care in England and the Information Services Division in Scotland. All participants gave their written consent to take part in the study.

6.2.2 Data Definitions

Menopause

Women were classified as postmenopausal if they reported their periods stopped naturally or after bilateral oophorectomy, or if they were aged 53 years or older (most women aged 53 years who had a natural menopause were postmenopausal by that age).

Parity and breastfeeding

At recruitment, women were asked to report the number of children they had given birth to (including still births), which was used to define parity. Women were asked to report, for each birth, if they had breastfed and if so, the duration of breastfeeding in months. This information was used to define breastfeeding (ever or never), and total duration of breastfeeding (summation over all children of reported duration of breastfeeding in months). In addition to using women's total duration of breastfeeding I also calculated the average duration of breastfeeding per child using each woman's total duration of breastfeeding divided by her parity. No questions were asked about the exclusivity of breastfeeding.

Ischaemic heart disease (IHD)

The outcome assessed in these analyses was first hospital admission (primary diagnosis) or death (primary underlying cause of death) due to ischaemic heart disease (IHD, ICD-10 I20-25). [20]

Co-factors & effect modifiers

Using data from the recruitment questionnaire participants in the Million Women Study were classified into 10 geographic regions of residence within England and Scotland. Women's socioeconomic status was derived using the Townsend deprivation index and summarised into quintiles. [22] Body mass index (BMI) was calculated as a woman's weight in kilograms divided by her height in meters squared and summarised into categories as follows <22.5, 22.5 to 24.9, 25 to 27.4, 27.5 to 29.9, 30 to 32.4, 32.5 to 35, or $\geq 35 \text{ kg/m}^2$ (I used 22.5 to 24.9 kg/m^2 as the reference category). Women were categorised as never, past, or current smokers. Current smokers were further categorised by the number of cigarettes smoked per day as follows, 1 to 4, 5 to 9, 10 to 14, 15 to 19, 20 to 24, or ≥ 25 . Women's alcohol intake was summarised using information on the number of glasses of wine, half pints of beer/lager, or measures of spirits drunk in an average week. This information was converted into standard number of drinks consumed per week and alcohol intake was categorised as none, 1 to 6, 7 to 14, or ≥ 15 units per week. Strenuous physical activity was summarised into four categories, none, ≤ 1 time per week, 2 to 3 times per week, ≥ 4 times per week. Information on co-factors was taken as reported at recruitment.

Where appropriate I adjusted for measurement error in BMI, smoking, and alcohol using the regression dilution approach; [23, 24] I summarised each of the co-factors into categories based on information from the recruitment questionnaire, however the average for each category was calculated for smoking and alcohol using data as reported on the repeat survey questionnaire (administered about 3 years after recruitment) from 708,265 women, and for BMI using data from a random sample of 3,745 women with measured data taken about 8 years after recruitment. [25]

6.2.3 Statistical Analysis

Eligibility and exclusion criteria

Women were eligible for inclusion into the main analysis if they were postmenopausal, did not have a registered diagnosis of cancer (except non-melanoma skin cancer) at recruitment, and had been asked about their breastfeeding histories. Of the 979,147 women who were eligible women missing information on parity (0.2%), or breastfeeding (11%) were excluded. For the main analysis women with a hospital admission for an IHD event (any diagnosis) before study entry (0.7%) were also excluded. The final number of women included in the main analysis was 862,861 (88% of those eligible). Women missing values for any of the adjustment variables were assigned to a separate category for that variable. Figure 6.1 (page 191) shows a flow diagram of the eligibility and inclusion criteria and the numbers of women excluded.

Exposures, outcome, co-factors, & effect modifiers

The outcome assessed in these analyses was first hospital admission or death due to IHD. The main exposure variables were parity and duration of breastfeeding treated as categorical variables. Parity was summarised as 0, 1, 2, 3, or 4 or more; and total duration of breastfeeding as never breastfed, breastfed for less than 6 months, breastfed for 6 to 9 months, or breastfed for 10 months or more (these approximate quartiles of total duration for women who breastfed).

All analyses were region of residence (10 regions) and adjusted for age (5 year categories), socioeconomic group (in fifths), and fine categories of smoking, alcohol, and

strenuous physical activity. The potential modifying effect of BMI on any associations between childbearing and women's risk of IHD was also explored.

Model specifics

I used Cox proportional hazards models in STATA version 11 (Stata corporation, TX, USA) to obtain hazard ratios (HR) and corresponding 95% confidence intervals (CI) of IHD associated with women's childbearing histories; I estimated the effects of parity and breastfeeding separately and in combination. I used attained age as the underlying time variable. Women contributed person-years from recruitment until the date of their first IHD event, death, emigration or end of follow-up, whichever was earliest. Hospital and death data are complete until 31/3/2008 for women recruited in England and 31/12/2008 for women recruited in Scotland. When comparing risks across more than two categories I estimated variances by treating the HR as floating absolute risks. [26] Where no natural baseline existed I used the category containing the largest number of women as the reference. Women of parity two (mean parity in this study population was 2.1 children [SD 1.0]) and who did not breastfeed were the baseline group I used to calculate the incidence rates for IHD per 1000 women over 5 years.

Additional analyses

I examined for modification of the effects of parity or breastfeeding by other factors reported at recruitment and known to affect IHD risk, for example prevalent diabetes, hypertension and high total blood cholesterol. I also examined for effect modification by a history of hypertension during pregnancy. I performed sensitivity analyses restricting our analyses to women with complete information for all relevant factors.

I also repeated our analyses including the 7,424 women considered ineligible for the main analysis. To explore the effects of sub-clinical disease on our model estimates I repeated the main analysis excluding women with an event during the first three years of follow-up.

6.3 Results

Characteristics of the study population

The main analyses included 869, 282 postmenopausal women. Figure 6.1 (page 191) describes the eligibility and exclusion criteria applied to women in the study. There were 35,590 incident events during 7,735,000 person-years of follow-up time. The average incidence rate was 4.6 per 1000 women years (CI 4.5 to 4.7). Table 6.1 (page 192) shows the baseline population characteristics for women included in the main analysis. Women were aged 57.5 years [SD 4] on average at study entry and had a mean BMI of 26.2 kg/m² [SD 5]. 88% were parous, with mean parity of 2.1 children [SD 1.2]. Among parous women, 60% had ever breastfed and the average total duration of breastfeeding was 4.7 months [SD 7.8]. The mean age of first birth was 23.9 years (SD 4.2).

The proportion of women reporting ever breastfeeding increased with increasing parity as did women's total duration and per child duration of breastfeeding. Compared to parous women, nulliparous women were less deprived, less likely to be current smokers, more likely to reported engaging in strenuous physical activity, consumed more alcohol, had a lower mean BMI than parous women (25.8 kg/m²) and had the lowest crude rate of incident IHD (4 per 1000 women per year, CI 3.8 to 4.1). Among parous women, those of parity three or more were more deprived, more likely to smoke, less

likely to report engaging in strenuous physical activity, drank less, had a higher mean BMI (26.7 kg/m^2), and a higher crude rate of IHD (5.7 per 1000 women per year, CI 5.6 to 5.8) than women who had fewer births. Parous women who reported ever breastfeeding were less deprived, less likely to be current smokers, were more likely to exercise, consumed slightly more alcohol, had a lower mean BMI (26.1 kg/m^2), and had a lower crude rate of IHD (4.6, CI 4.5 to 4.7) than women who reported never breastfeeding (4.9 CI 4.8 to 5.0).

Univariate analyses

Table 6.2 (page 193) shows the HR of an IHD event in models adjusted for age and region for parity, breastfeeding and selected co-factors. These minimally adjusted models explore the individual effects of each variable on the risk of IHD. A history of having children was associated with a modest increase in the risk of IHD (HR 1.30, CI 1.27 to 1.38). Among parous women those with a history of ever breastfeeding had a lower risk of IHD than parous women who did not breastfeed (HR 0.86, CI 0.84 to 0.89). Smoking and then BMI had the largest effects on women's risk of IHD; women who reported currently smoking 15 or more cigarettes per day had more than double the risk of an IHD event than women who reported being never smokers (HR 2.58, CI 2.39 - 2.80). In terms of BMI, women with a BMI over 30 kg/m^2 had a risk of IHD 77% greater than women with a BMI less than 25 kg/m^2 (HR 1.77, CI 1.65 to 1.90). The effects of having children and of breastfeeding were similar in size to the effects of changing socioeconomic group or being physically inactive.

Figures 6.2 - 6.7 (pages 194 - 196) show the incidence rate for an IHD event per 1000 women over 5 years for parity, duration of breastfeeding (limited to parous women),

smoking, BMI, socioeconomic group, and levels of strenuous physical activity. The strong association between an increasing number of cigarettes smoked and the incidence of IHD is clear (figure 6.4 , page 195). Compared to women who reported never smoking, women who reported smoking 25 or more cigarettes per day had an excess of 40 IHD events per 1000 women over 5 years. The association between increasing BMI and increasing incidence of IHD in this population can also be seen; compared to women with a BMI $<25 \text{ kg/m}^2$ women with a BMI between 25 to 30 kg/m^2 or $\geq 30 \text{ kg/m}^2$ had an excess of about 25 and 34 IHD events per 1000 women over 5 years respectively. In comparison, for parity and breastfeeding: compared to parity of two women of parity ≥ 3 had an excess of about 10 IHD events per 1000 women over 5 years; compared to parous women who did not breastfeed women who breastfed for ≥ 10 months had 3 fewer IHD events per 1000 women over 5 years.

Multivariate analyses

Table 6.3 (page 197) shows the IHD hazard ratios in women by parity and for parous women by breastfeeding status in models adjusted for selected co-factors, and in the final column adjusted for age, region, socioeconomic group, BMI, smoking, alcohol, strenuous physical activity, and parity and breastfeeding as appropriate. Women of parity two who did not breastfeed were used as the reference category. Table 6.4 (page 198) shows the χ^2 statistic for heterogeneity across categories of parity (three degrees of freedom) and duration of breastfeeding (three degrees of freedom).

Comparing nulliparous to parous women

In the model adjusted only for age, socioeconomic group, and region parous women had a 30% higher risk of an incident IHD than nulliparous women. Adjusting for either BMI or smoking slightly weakened the association; however, adjusting for both lowered the risk further. Adding additional adjustments for alcohol and strenuous physical activity had little effect on estimated risk of IHD associated with being parous, after full adjustment nulliparous women still had about a 20% lower risk of IHD than parous women (HR 1.18, CI 1.13 to 1.23).

Risk associated with increasing parity

In a model adjusted for age, socioeconomic group, and breastfeeding, and region, increasing parity was associated with an excess risk of incident IHD. Again adding adjustment for both smoking and BMI weakened the effect more than adjusting for each co-factor alone. All co-factors included in the final model were statistically important. In the final model, parity above two (i.e. ≥ 3) was still associated with a small increase in the risk of IHD. For all the models, tests for linear trend were significant ($p < 0.0001$). In the final model, for women with a parity ≥ 3 , an increasing parity was associated with a hazard ratio of 1.09 (CI 1.09 to 1.10) per one unit increase in parity.

Comparing women who breastfed to women who did not breastfeed

Among parous women a history of having ever breastfed was associated with a 14% decrease in the risk of an incident event (HR 0.86, CI 0.83 to 0.88) in models adjusted only for age, region, and socioeconomic group. Again, adding adjustment for both

smoking and BMI weakened the association. In the final model (which included additional adjustment for alcohol and strenuous physical activity) parous women who ever breastfed had an 8% lower risk of an IHD event than parous women who never breastfed risk (HR 0.92, CI 0.90 to 0.95).

Risk among parous women by duration of breastfeeding

In a model adjusted only for age, socioeconomic group, and region, an increasing duration of breastfeeding was associated with a cumulative decrease in the risk of IHD. Table 6.3 (page 197) also shows the lengths of breastfeeding duration captured in these categories and the mean duration of breastfeeding within in each tertile. It can be seen that the durations of breastfeeding are unequal and that the mean total duration of breastfeeding increased exponentially across the categories.

Adding important co-factors to the model weakened the association between duration of breastfeeding and risk of IHD (more so with longer durations of breastfeeding). In the fully adjusted model, women who breastfed for an average of 2 months had a very similar risk (HR 0.95, CI 0.93 to 0.97) to women who breastfed on average 7 months (HR 0.93, CI 0.90 to 0.97). In comparison, women who breastfed for ≥ 10 months (19 months on average) had a much lower risk (HR 0.85, CI 0.83 - 0.89). For all models preceding the main model, tests for linear trend for breastfeeding duration in months were significant ($p < 0.0001$). In the main model there was no statistical evidence of a linear trend ($p = 0.14$). When a Likelihood Ratio-test for trend associated with increasing durations of breastfeeding was restricted to women who had breastfed the trend was borderline significant ($p = 0.06$); when duration of breastfeeding was modeled as a continuous variable there was a small statistically significant decrease ($p < 0.0001$) in

the risk of IHD associated with longer durations of breastfeeding (HR 0.97 for every 6 months of breastfeeding, CI 0.96 to 0.98).

Findings in context

The contributions of BMI and smoking to the risk of an IHD event

As can be seen by the results in table 6.3 (page 197) smoking and BMI strongly affect the risk of IHD associated with childbearing. Additional evidence of this can be seen by comparing the contribution of selected co-factors to size of the associated the χ^2 statistic for heterogeneity. Table 6.4 (page 198) shows the values of the χ^2 statistic for heterogeneity across categories of parity and duration of breastfeeding on three degrees of freedom for each of the models (i.e. adjusted for smoking, BMI, and other co-factors) in table 6.3 (page 197). For parity the biggest change in the size of the associated χ^2 statistic were with the additions of BMI and smoking to the model. The contribution of BMI and smoking were similar, adjusting for both factors reduced the associated χ^2 statistic by nearly half (from 676.12 to 360.73). In contrast, for breastfeeding smoking had a bigger effect on the χ^2 statistic than BMI; adding smoking to the model had the biggest single effect on the size of the associated χ^2 statistic reducing it by nearly half (from 283.28 to 170.88). For both parity and breastfeeding there is a large difference in the size of the associated χ^2 statistics between minimally adjusted and fully adjusted models. After full adjustment the χ^2 statistic for heterogeneity across categories of parity was reduced by 66%, for categories of breastfeeding duration it was reduced by 75%.

The contributions of childbearing and other co-factors to the risk of an IHD event

Table 6.5 (page 202) shows the IHD hazard ratios obtained from the final model (adjusted for age, socioeconomic group, BMI, smoking, alcohol, strenuous physical activity, region, parity, and duration of breastfeeding as appropriate) for selected variables included in the final model. As in the univariate analyses smoking and BMI were strongly associated with risk of IHD; the relative effect size of being a smoker or of having a raised BMI was much greater than the effect associated with a change in exposure level for parity, breastfeeding, socioeconomic group, or strenuous physical activity.

Figures 6.8 -6.13 (pages 199 - 201) show the incidence rates of an IHD event for childbearing and selected co-factors (per 1000 women over 5 years). The adjusted rates were calculated using women of parity two who had never breastfed as the baseline rate and multiplying this by the hazard ratios obtained from the fully adjusted model in table 6.3 (page 197). The scale of the y-axis on each of on each of these figures is the same. Smoking and BMI continued to have important effects on women's risk of an IHD event after the menopause. Being physically inactive or having a parity ≥ 3 was associated with a similar excess in the risk of an IHD event. Duration of breastfeeding was weakly associated with women's risk of an IHD event.

Additional analyses**Effect on model estimates of altering the study population**

Table 6.6 (page 203) shows the effects on model estimates of including women considered ineligible for the main analysis (7,424 excluded because of a prior hospital admission for an IHD event), of restricting the study population to women with complete

information for all factors included in the model, or of excluding women with an IHD event during the first three years of follow-up. Altering the study population by these methods did not materially effect the final model estimates for the risk of incident IHD associated with childbearing.

Effects on model estimates of other co-factors which may influence women's risk of IHD

Table 6.7 (page 204) shows the effects on the IHD hazard ratio obtained in the final model of including adjustments for self-reported diabetes, high blood pressure or high blood cholesterol, high blood pressure during pregnancy, or ethnicity which were not included in the final model used in the main analysis to avoid over-fitting. Adjusting for each of these factors did not materially affect the results.

6.4 Discussion

Summary

In this analysis which included more than 800,000 postmenopausal women aged 57.6 years on average, with a mean BMI of 26.2 kg/m^2 a parity ≥ 3 was associated with a small increase in women's risk of IHD after the menopause. For parous women, a history of breastfeeding was associated with a small decrease in the risk and there was a small but statistically significant trend associated with increasing durations of breastfeeding. The size of the effect on risk of IHD for increasing parity was similar to the size of the effects associated with moving from the middle to the lowest socioeconomic

group or of becoming physically inactive. These effects were small in comparison to the excess risk of IHD associated with current smoking and with increasing BMI.

Findings in relation to other studies

The age-standardised incidence rates for IHD found in this study were comparable to other population-based data from the UK. [27, 28] The association between childbearing and excess risk of cardiovascular disease has been previously reported. [7, 9, 10] However in contrast to the only other population-based prospective analysis, carried out in Swedish women of a similar age to Million Women Study participants, I did not find a "J-shaped" association between parity and risk of incident IHD. [12] Among Million Women Study participants nulliparity was associated with a lower risk of IHD compared to parous women.

There are only two other prospective studies that have explored the association between breastfeeding and women's later risk of heart disease, and they reported conflicting findings.[13, 14] Schwartz et al found that duration of breastfeeding had no effect on the incidence of cardiovascular disease in a population of 139,681¹ postmenopausal women after 8 years of follow-up.[13] In contrast Stuebe et al reported finding a lower risk of incident myocardial infarction associated with breastfeeding in a population of 89,362 (2,540 incident events) middle-aged women (and with adjustment made for menopausal status); however, in this study the lower risk was limited to women with the longest durations of breastfeeding (23 or more months), which made up less than 6% of the study population.[14] The study population used in this analysis is substantially bigger than either of these (869,282 women and 35,590 incident events). The risk estimates reported were similar to those found in both studies, and in common

¹Number of events not reported

with both studies the biggest reduction in risk was limited to the small group women with the longest durations of breastfeeding; HR 0.86, CI 0.38 to 0.89 for parous Million Women Study participants (total breastfeeding duration > 10 months), HR 0.77, CI 0.62 – 0.94 for parous Nurses Health Study participants (total breastfeeding duration > 23 months), and HR 0.93, CI 0.85-1.02 for Women's Health Initiative participants (total breastfeeding duration > 24 months). As compared to the number of IHD events in Million Study participants the reference category (parity two, never breastfed) these relative risks would be associated with roughly 3, 5, or 1 fewer IHD events per 1000 women over 5 years.

Possible explanations

Several co-factors included in the main analysis, for example BMI, are known to be imperfect measures. [29, 30] It was clear in all of the analyses that smoking and BMI were strongly associated with women's risk of IHD, findings which are consistent with the published literature. [6, 31] Adjusting for these co-factors weakened the associations between parity and breastfeeding and IHD risk, and reduced the size of the associated χ^2 statistics. It is reasonable to expect that if it were possible to include better measures of these variables the value of χ^2 statistic would be reduced further. In the final model the size of the association between parity more than two and risk of IHD was similar in magnitude to that of changes in socioeconomic group. However, it is known that the excess risk of disease associated with socioeconomic status is largely as a result of differences in factors such as smoking and BMI which again highlights the importance of these risk factors which are modifiable even in older population groups. [4, 32] As it is not possible to eliminate residual confounding due to imperfectly measured variables

it is prudent to interpret the small independent effects of childbearing on women's risk of IHD after the menopause with caution.

In terms of breastfeeding the size and nature of the association between longer durations of breastfeeding and risk of IHD changed with adjustment for important co-factors. In comparing findings from the Million Women Study with the published findings from the Women's Health Initiative, and Nurses Health Study, for all three studies the lower risk of IHD associated with duration of breastfeeding was primarily driven by the very low risk among women with the longest duration of breastfeeding category (≥ 10 months among women in the Million Women Study, > 24 months in the Womens Health Initiative and > 23 months in the Nurses Health Studies). There is historical evidence that rates of breastfeeding were low and durations were short during the 1960s and 1970s which is when the majority of women in these three studies would have been having children (see 3, page 63 for detailed review and analysis of historical trends in breastfeeding). In each of these studies, these women made up a small proportion of the total study population (less than one fifth) and were quite different in their baseline characteristics to women with shorter durations of breastfeeding. For example, among Million Women Study participants women who breastfed the longest were much less likely to be current smokers (half as likely to report being current smokers at study entry). In the current analysis being a current smoker was associated with an excess risk of IHD that was more than twice that of non-smokers (HR 2.07, CI 2.05 to 2.17). In comparison, the longest duration of breastfeeding was associated with a 14% risk reduction (HR 0.86, CI 0.83 to 0.89), a decrease that is within this error margin.

Thus, although my findings are similar to those reported in the literature, and although I found statistical evidence that childbearing influences women's risk of IHD, I was not able to exclude residual confounding as the cause of any remaining association between

women's childbearing history and risk of IHD after the menopause. It is possible that breastfeeding has an indirect effect on women's risk of IHD mediated through the effects of a lower BMI on blood pressure, total cholesterol, and risk of diabetes mellitus.

Strengths and weaknesses

Strengths of this study include its size and the availability of information about co-factors. The Million Women Study includes one in four UK women aged between 50 and 64 years at the time of recruitment. This analysis included 869,282 postmenopausal women and over 35,000 cases of ischaemic heart disease which to my knowledge is the sample-size to date. Detailed information on socioeconomic factors, smoking and physical activity allowed for fine subdividing by key co-factors whilst retaining sufficient power to assess the effects of childbearing and breastfeeding on IHD risk. Updated exposure information was also available which allowed me to correct for possible misclassification errors. [23, 24] The study population is ethnically homogeneous with 98% reporting they were of white ethnicity.

As with most large-scale epidemiological studies exposure information was obtained by self-report. Studies in older women have found that their long-term recall of parity and breastfeeding is reliable, [33, 34] and I have shown previously that self-reported breastfeeding data corresponds well to data obtained from historical data sets (Chapter 3, page 63). Given the historical trends in breastfeeding there is substantial power to assess the effects of short durations of breastfeeding but less power to assess the effects of long durations of breastfeeding if they exist.

For this analysis I used outcome data from hospital episode statistics and death certificates. In the UK death certification data are reliable and complete. [15] Data from

hospital admissions has been shown to be reasonably reliable. [35–37] As the hospital admissions databases started in from January 1981 in Scotland and April 1997 in England it is possible that a woman's first admission with IHD during the Million Women Study follow-up period was not her first IHD event.[18, 19] If a true effect of childbearing existed on a woman's risk of a first ischaemic event it could be masked by the increase in risk of a second IHD event (56,000 women in this study population reported a diagnosis of or treatment for heart disease at recruitment), however again any effect would likely still be small and overshadowed by known risk factors.

Conclusions

For women in the UK there is evidence that childbearing is modestly associated with the risk of ischaemic heart disease after the menopause. Importantly though, known risk factors that women can modify now, for example smoking less or stopping completely and losing weight, are more important in mediating this risk than past childbearing.

6.5 Chapter summary

In this chapter I explored the associations between women's childbearing histories and their risk of an incident IHD event. I found that increasing parity was still statistically and meaningfully associated with a higher risk of IHD after adjusting for important co-factors. I also found that increasing duration of breastfeeding was statistically associated with a lower risk, however the reduction in number of events is very small. In keeping with findings presented in chapter 5 (page 122) I found that childbearing had a

limited association with risk of IHD compared to other known risk factors. In contrast, I found the associations were independent of BMI.

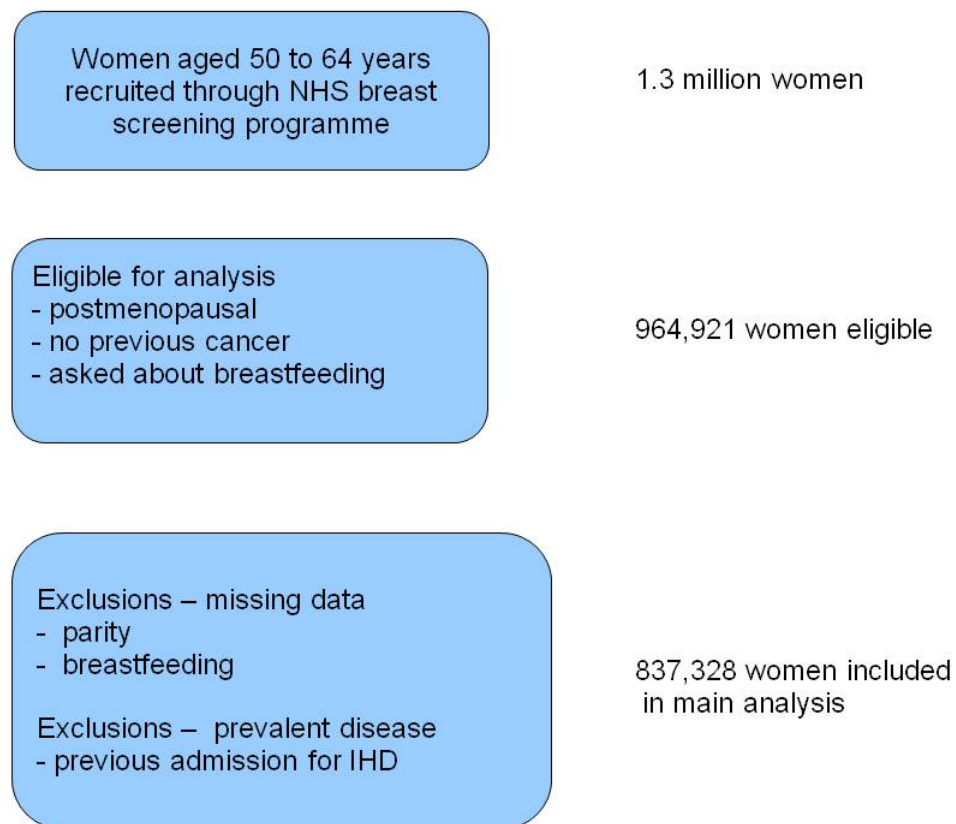


FIGURE 6.1: Flow diagram showing eligibility and inclusion criteria and the final number of women included in the main analyses

TABLE 6.1: Characteristics of study participants included in the main analyses of the effects of childbearing on women's risk of ischaemic heart disease according to parity and breastfeeding status

Population Characteristics	By Parity		By Breastfeeding Status*		All Women
	Nulliparous n = 101 331	One or two n = 462 868	Never Breastfed n = 228 945	Ever Breastfed n = 507 052	n = 837 328
Mean age in years (SD)	57.5 (5)	57.3 (4)	56.7 (4)	58 (4)	57.5 (4)
Mean parity (SD)	0	1.7 (0.4)	2.1 (1)	2.5 (1)	2.1 (1)
Ever breastfed (%)	0	64	0	100	69
Mean number children breastfed (SD)	0	1 (0.8)	0	2.0 (1.0)	1.3 (1.2)
Duration breastfeeding					
Total lifetime duration in months (SD)	0	3.7 (5.5)	0	7.7 (8.9)	4.7 (7.9)
Average duration per child in months (SD)	0	2.1 (3.0)	0	3.1 (3.2)	1.9 (2.9)
In lowest socioeconomic tertile (%)	32	30	38	32	33
Current smokers (%)	16	17	22	16	18
Engaging in strenuous physical activity two or more times per week (%)	21	20	17	21	20
Average alcohol consumption (g/day) (SD)	6.3 (8.3)	5.8 (7.4)	5.3 (7.1)	5.8 (7.4)	5.7 (7.4)
Mean body mass index (SD)	25.8 (5)	26.0 (4.5)	26.5 (5)	26.2 (4)	26.3 (5)
Follow Up					
Person-years of follow-up (1000)	902	4140	2026	4539	7467
Number of women with first admission or death due to IHD	3787	17792	10396	21820	36003
IHD incidence rates per 1000 (95%CI)	4.2 (4.1 to 4.3)	4.3 (4.2 to 4.4)	5.1 (5.0 to 5.2)	4.8 (4.7 to 4.9)	4.8 (4.8 to 4.9)

* Amongst parous women only
Percentages exclude the small number of women with missing values

TABLE 6.2: Risk of an IHD event for selected characteristics in models adjusted only for age and region (HR, 95%CI)

Characteristics of women		Cases	Person years at risk (1000)	IHD Rate (per 5000 women-years)	(95% CI)	HR	(95% CI)
Parity							
	Nulliparous	3787	902	4.20	(4.06 - 4.33)	1.00	-
	Parous	32,216	6600	4.91	(4.85 - 4.96)	1.28	(1.23 - 1.33)
LR test heterogeneity							p < 0.0001
Breastfed							
	No	10,396	2000	5.13	(5.03 - 5.23)	1.00	-
	Yes	21,820	4500	4.81	(4.74 - 4.87)	0.85	(0.84 - 0.87)
LR test heterogeneity							p < 0.0001
Socioeconomic group							
	Upper third	8978	2500	3.62	(3.55 - 3.69)	0.82	(0.75 - 0.90)
	Middle third	11,165	2500	4.42	(4.34 - 4.50)	1.00	(0.91 - 1.09)
	Lowest third	15,860	2500	6.44	(6.34 - 6.54)	1.43	(1.30 - 1.56)
Body mass index							
	< 20	1086	284	3.82	(3.60 - 4.05)	0.99	(0.94 - 1.07)
	20 to 22.4	3871	1200	3.32	(3.22 - 3.42)	0.87	(0.83 - 0.90)
	22.5 to 24.9	7306	1900	3.87	(3.79 - 3.96)	1.00	(0.97 - 1.03)
	25 to 27.4	7762	1600	4.89	(4.78 - 5.00)	1.23	(1.19 - 1.28)
	27.5 to 29.9	5381	961	5.60	(5.45 - 5.75)	1.41	(1.35 - 1.46)
	30 to 32.4	3662	572	6.39	(6.19 - 6.60)	1.61	(1.55 - 1.68)
	32.5 to 34.9	2162	312	6.91	(6.63 - 7.21)	1.77	(1.68 - 1.85)
	≥ 35	2736	335	8.16	(7.86 - 8.47)	2.15	(2.06 - 2.25)
LR test heterogeneity							p < 0.0001
Smoking							
	Never	13,467	3700	3.62	(3.56 - 3.68)	1.00	(0.93 - 1.08)
	Past	10,528	2000	5.21	(5.11 - 5.31)	1.45	(1.35 - 1.57)
	Current <15	4796	686	6.99	(6.79 - 7.19)	2.09	(1.93 - 2.26)
	Current ≥15	5245	626	8.37	(8.15 - 8.60)	2.67	(2.46 - 2.89)
LR test heterogeneity							p < 0.0001
Strenuous physical activity							
	< 1 per week	28,975	5700	5.08	(5.02 - 5.14)	1.00	(0.90 - 1.11)
	≥ 1 per week	5344	1500	3.56	(3.45 - 3.64)	0.72	(0.65 - 0.81)
LR test heterogeneity							p < 0.0001

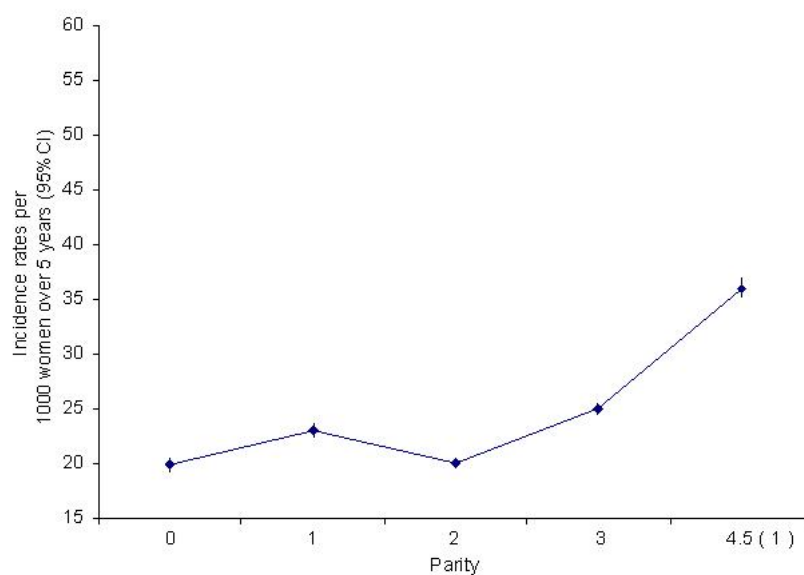


FIGURE 6.2: Incidence rate for IHD per 1000 women over a 5 year period by parity

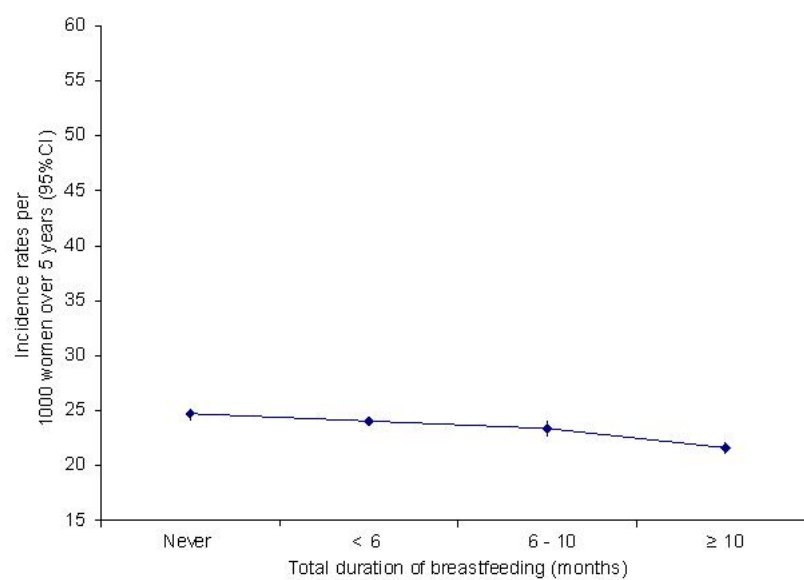


FIGURE 6.3: Incidence rate for IHD per 1000 women over a 5 year period among parous women by duration of breastfeeding

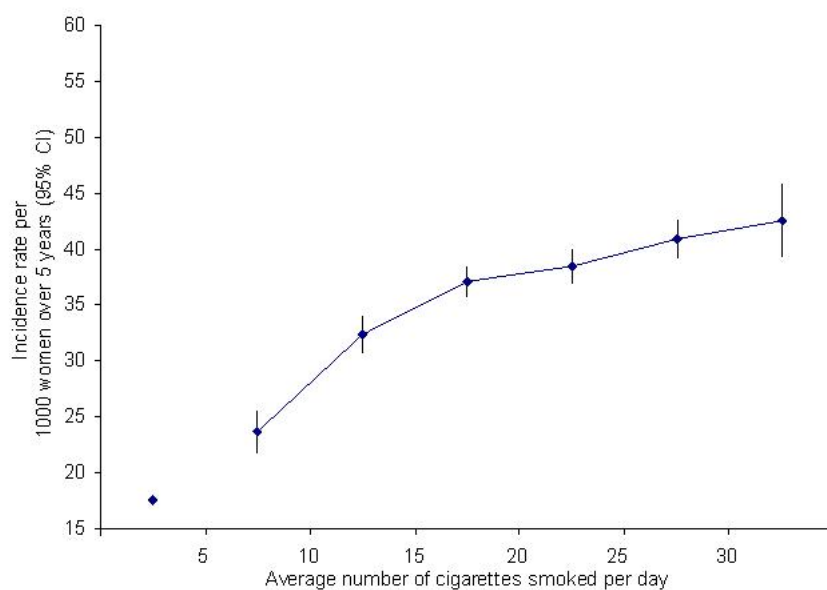


FIGURE 6.4: Incidence rate for IHD per 1000 women over a 5 year period by average number of cigarettes smoked per day

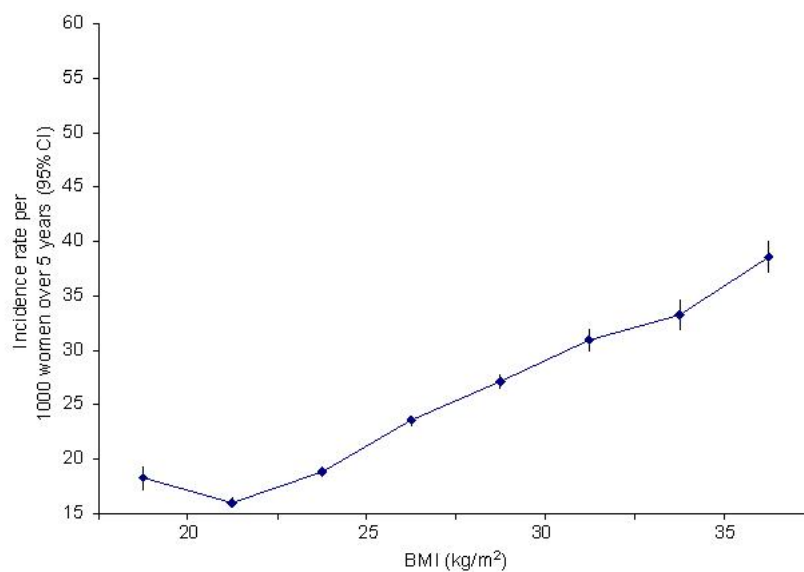


FIGURE 6.5: Incidence rate for IHD per 1000 women over a 5 year period by BMI

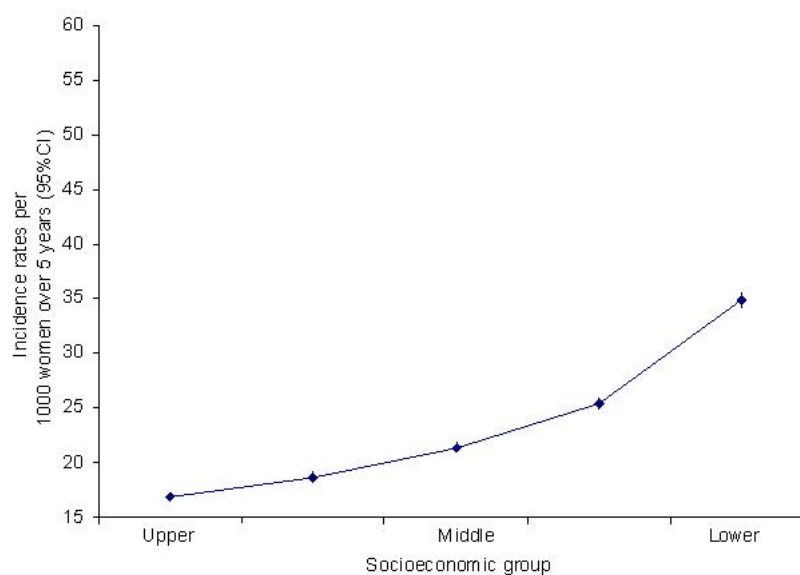


FIGURE 6.6: Incidence rate for IHD per 1000 women over a 5 year period by socioeconomic group

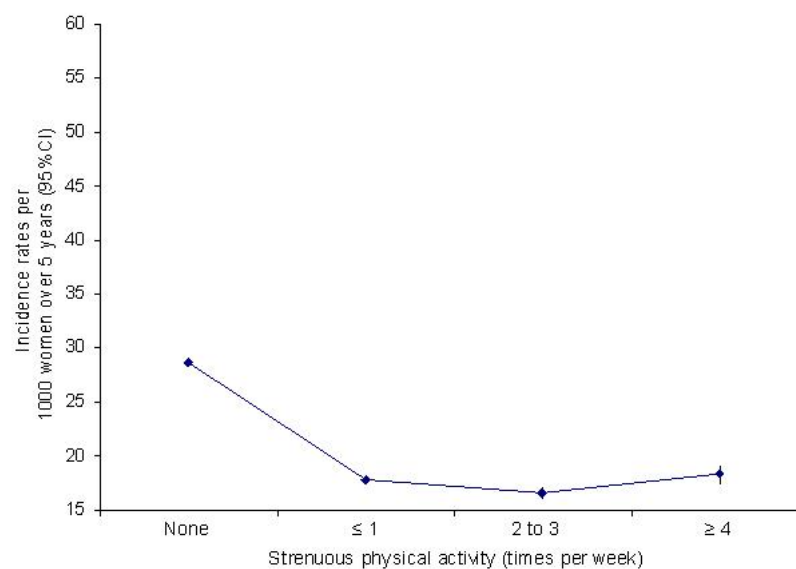


FIGURE 6.7: Incidence rate for IHD per 1000 women over a 5 year period by amount of strenuous physical activity per week

TABLE 6.4: χ^2 statistic for heterogeneity across categories of parity (three degrees of freedom) and durations of breastfeeding (three degrees of freedom) for each of the models in table 6.3 (page 197)

Chi-square for heterogeneity across categories		
	Parity [3 degrees of freedom]	Duration of breastfeeding [3 degrees of freedom]
Model A - Adjusted for age, socioeconomic group, parity, and breastfeeding as appropriate, and stratified by region	681.67	315.5
Model A + additionally adjusted for BMI	512.25	263.42
Model A + additionally adjusted for smoking	511.35	193.74
Model A + additionally adjusted for BMI and smoking	352.87	147.74
Model A + additionally adjusted BMI, smoking, alcohol, and strenuous physical activity	280.22	83.71

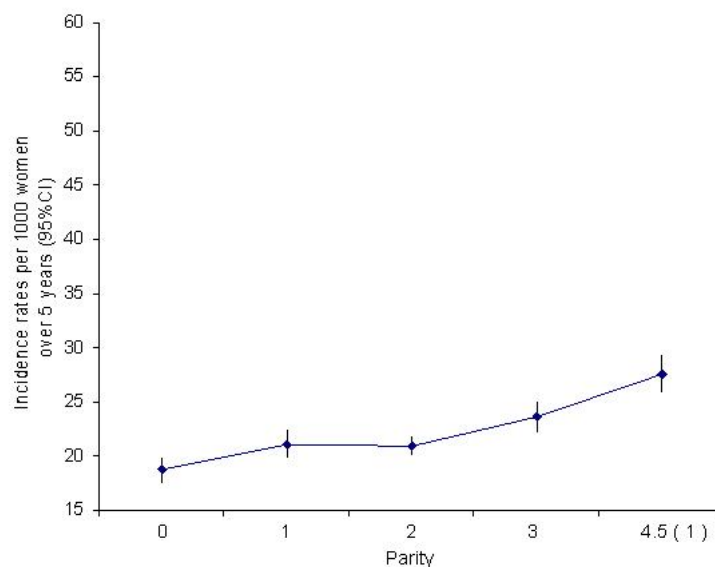


FIGURE 6.8: Incidence rate for ischaemic heart disease per 1000 women over a 5 year period among parous women by parity with adjustment for age, socioeconomic group, smoking, BMI, alcohol, strenuous physical activity, and duration of breastfeeding as appropriate, and region

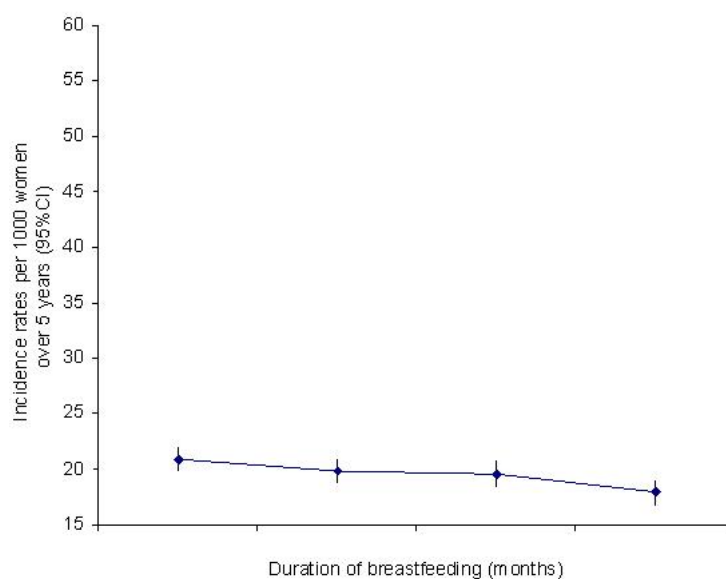


FIGURE 6.9: Incidence rate for ischaemic heart disease per 1000 women over a 5 year period among parous women by duration of breastfeeding with adjustment for age, socioeconomic group, smoking, BMI, alcohol, strenuous physical activity, and parity as appropriate, and region

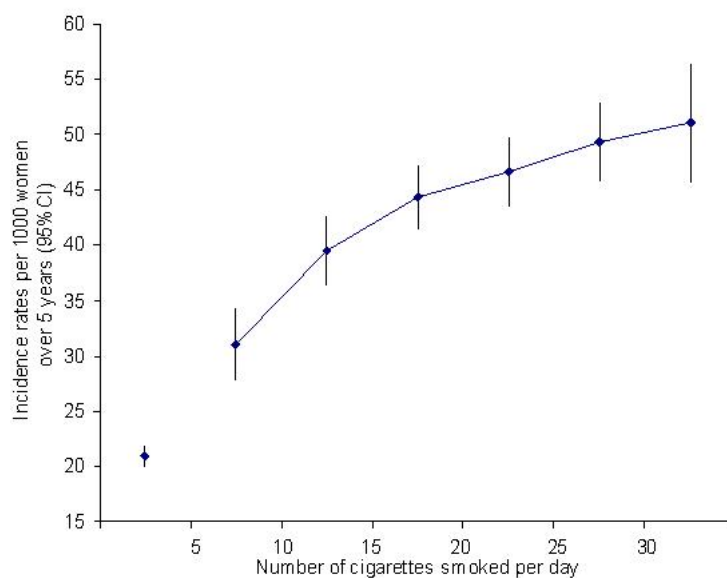


FIGURE 6.10: Incidence rate for ischaemic heart disease per 1000 women over a 5 year period among parous women by number of cigarettes smoked with adjustment for age, socioeconomic group, BMI, alcohol, strenuous physical activity, parity, and duration of breastfeeding as appropriate, and region

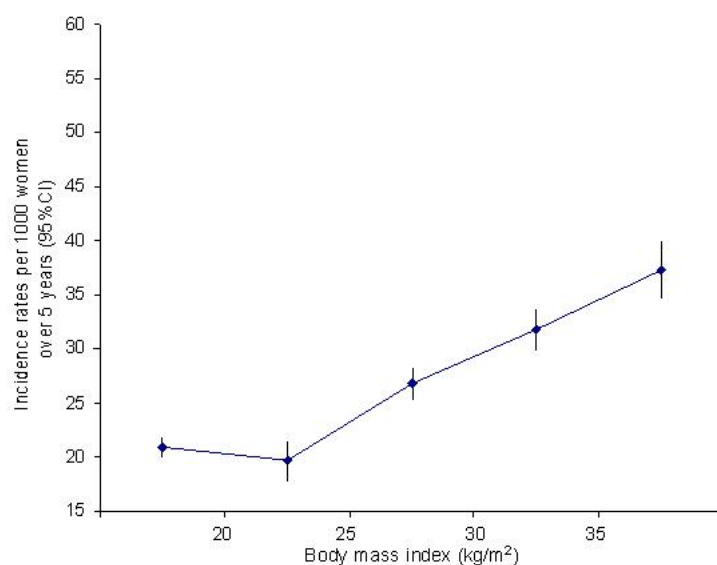


FIGURE 6.11: Incidence rate for ischaemic heart disease per 1000 women over a 5 year period among parous women by BMI with adjustment for age, socioeconomic group, smoking, alcohol, strenuous physical activity, parity, and duration of breastfeeding as appropriate, and region

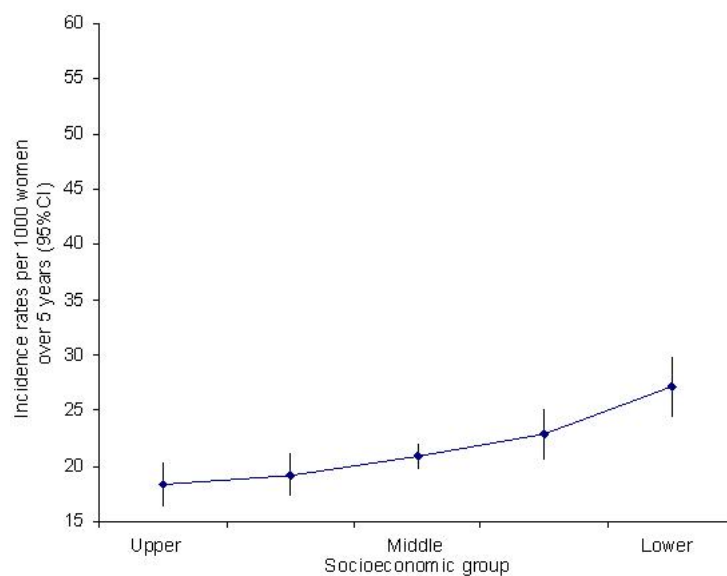


FIGURE 6.12: Incidence rate for ischaemic heart disease per 1000 women over a 5 year period among parous women by socioeconomic group with adjustment for age, smoking, BMI, alcohol, strenuous physical activity, parity, and duration of breastfeeding as appropriate, and region

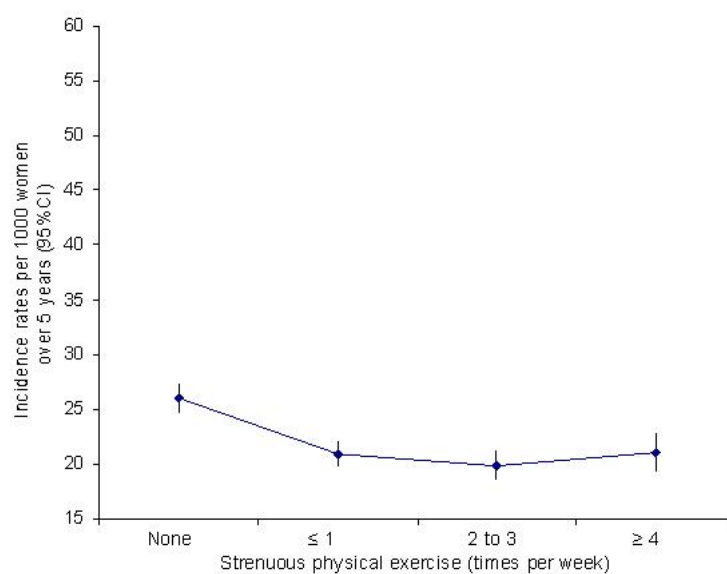


FIGURE 6.13: Incidence rate for ischaemic heart disease per 1000 women over a 5 year period among parous women by amount of strenuous physical activity per week with adjustment for age, socioeconomic group, smoking, BMI, alcohol, parity, and duration of breastfeeding as appropriate, and region

TABLE 6.5: Risk of ischaemic heart disease among parous women in a model adjusted for age, socioeconomic group, smoking, BMI, alcohol, strenuous physical activity, and region (HR,95%CI)

Characteristics of women		Cases	HR	(95% CI)
Parity				
	0	3,787	0.91	(0.87 - 0.94)
	1	4,904	1.02	(0.98 - 1.04)
	2	12,888	1.00	(0.99 - 1.01)
	3	8,304	1.13	(1.10 - 1.16)
	≥ 4	6,120	1.31	(1.27 - 1.35)
Breastfeeding duration				
	0	14,183	1.00	(0.99 - 1.01)
	< 6	11,975	0.94	(0.92 - 0.96)
	6 to 10	4,042	0.93	(0.90 - 0.97)
	≥ 10	5,803	0.85	(0.82 - 0.88)
Socioeconomic group				
	Upper	5,331	0.88	(0.81 - 0.94)
		5,948	0.92	(0.85 - 0.99)
	Middle	6,689	1.00	(0.99 - 1.01)
		7,787	1.10	(1.02 - 1.18)
	Lowest	9,994	1.30	(1.21 - 1.40)
Body mass index				
	< 20	1,086	0.94	(0.88 - 1.00)
	20 to 24.9	11,177	1.00	(0.99 - 1.01)
	25 to 29.9	13,143	1.28	(1.25 - 1.31)
	30 to 34.9	5,824	1.55	(1.51 - 1.60)
	≥ 35	2,736	1.88	(1.80 - 1.96)
Smoking				
	Never	13,467	1.00	(0.98 - 1.02)
	Current <5	590	1.48	(1.36 - 1.60)
	Current 5 to 15	4,206	2.02	(1.96 - 2.08)
	Current 15 to 25	4,609	2.35	(2.27 - 2.42)
	Current - ≥ 25	636	2.54	(2.34 - 2.42)
Strenuous physical activity				
	None	18,591	1.24	(1.19 - 1.30)
	1 or less per week	7,361	1.00	(0.99 - 1.01)
	2 to 3 times	3,286	0.95	(0.90 - 1.01)
	4 or more times	1,449	1.01	(0.95 - 1.08)

TABLE 6.6: The effect on model estimates of including women who had had a hospital admission for IHD prior to study recruitment, of only including women with complete information on all relevant factors, and of excluding the initial 3 years of follow-up (HR, 95%CI)

	Final model** used in the main analyses [Model B]				Model B + women with a hospital record of IHD before recruitment				Model B - restricted to women with complete information for potential confounders				Model B - excluding the first three years of follow up			
	Mean	Cases	HR	95%CI	Cases	HR	95%CI	Cases	HR	95%CI	Cases	HR	95%CI	Cases	HR	95%CI
Parity																
Nulliparous women	0	3,739	1.00	-	4,247	1.00	-	3,160	1.00	-	2,732	1.00	-	2,732	1.00	-
Parous women	2	32,190	1.18	(1.13 - 1.23)	37,193	1.18	(1.14 - 1.23)	27,252	1.18	(1.13 - 1.23)	22,993	1.16	(1.10 - 1.21)	22,993	1.16	(1.10 - 1.21)
Ever breastfed*																
No	0	10,349	1.00	(0.99 - 1.01)	12,111	1.00	(0.99 - 1.01)	8,743	1.00	(0.99 - 1.01)	7,313	1.00	(0.99 - 1.01)	7,313	1.00	(0.99 - 1.01)
Yes	4.9	21,841	0.92	(0.90 - 0.95)	25,082	0.93	(0.91 - 0.95)	18,509	0.92	(0.89 - 0.95)	15,680	0.92	(0.90 - 0.95)	15,680	0.92	(0.90 - 0.95)
Parity*																
1	1	4,853	1.01	(0.97 - 1.04)	5,590	1.01	(0.99 - 1.04)	4,068	1.00	(0.97 - 1.03)	3,488	1.01	(0.98 - 1.05)	3,488	1.01	(0.98 - 1.05)
2	2	12,928	1.00	(0.99 - 1.01)	14,795	1.00	(0.98 - 1.02)	11,133	1.00	(0.98 - 1.02)	9,330	1.00	(0.98 - 1.02)	9,330	1.00	(0.98 - 1.02)
3	3	8,312	1.12	(1.09 - 1.16)	9,610	1.13	(1.11 - 1.16)	7,059	1.12	(1.10 - 1.15)	5,980	1.13	(1.10 - 1.16)	5,980	1.13	(1.10 - 1.16)
>4	4.4	6,097	1.31	(1.27 - 1.35)	7,198	1.34	(1.31 - 1.37)	4,992	1.30	(1.26 - 1.33)	4,195	1.28	(1.24 - 1.32)	4,195	1.28	(1.24 - 1.32)
Total duration of breastfeeding*																
Did not breastfeed	0	10,349	1.00	(0.99 - 1.01)	12,111	1.00	(0.98 - 1.02)	8,743	1.00	(0.98 - 1.02)	7,313	1.00	(0.97 - 1.02)	7,313	1.00	(0.97 - 1.02)
< 6 months	2.3	11,987	0.95	(0.93 - 0.97)	13,860	0.96	(0.94 - 0.97)	10,219	0.94	(0.93 - 0.96)	8,602	0.95	(0.93 - 0.97)	8,602	0.95	(0.93 - 0.97)
6 - 9 months	7.2	4,036	0.94	(0.90 - 0.97)	4,593	0.94	(0.91 - 0.97)	3,423	0.93	(0.90 - 0.97)	2,892	0.93	(0.90 - 0.96)	2,892	0.93	(0.90 - 0.96)
10 + months	18.6	5,818	0.86	(0.83 - 0.89)	6,629	0.86	(0.84 - 0.88)	4,867	0.85	(0.83 - 0.88)	4,186	0.86	(0.83 - 0.89)	4,186	0.86	(0.83 - 0.89)

* Parous women only

** Adjusted for age, socioeconomic group, BMI, smoking, alcohol, strenuous physical activity parity, and breastfeeding as appropriate, and stratified by region

TABLE 6.7: The effect on model estimates of adding additional selected co-factors (HR, 95%CI)

	Mean	Cases	Model B + additional adjustment for diabetes	Model B + additional adjustment for high cholesterol or high blood pressure	Model B + additional adjustment for history of high blood pressure in pregnancy	Model B + additional adjustment for ethnicity
Parity						
Nulliparous women	0	3,739	1.00	1.00	-	1.00
Parous women	2	32,190	1.19 (1.14 - 1.23)	1.18 (1.14 - 1.23)	-	1.21 (1.14 - 1.27)
Ever breastfed*						
No	0	10,349	1.00 (0.99 - 1.01)	1.00 (0.99 - 1.01)	1.00 (0.99 - 1.01)	1.00 (0.99 - 1.01)
Yes	4.9	21,841	0.94 (0.92 - 0.96)	0.93 (0.91 - 0.96)	0.94 (0.91 - 0.96)	0.94 (0.92 - 0.96)
Parity*						
1	1	4,853	1.00 (0.98 - 1.04)	1.00 (0.97 - 1.03)	1.01 (0.98 - 1.04)	0.96 (0.92 - 1.00)
2	2	12,928	1.00 (0.98 - 1.02)	1.00 (0.98 - 1.02)	1.00 (0.98 - 1.02)	1.00 (0.98 - 1.02)
3	3	8,312	1.13 (1.11 - 1.15)	1.14 (1.11 - 1.16)	1.13 (1.11 - 1.15)	1.12 (1.10 - 1.14)
>4	4.4	6,097	1.31 (1.28 - 1.34)	1.33 (1.30 - 1.36)	1.32 (1.29 - 1.36)	1.31 (1.27 - 1.36)
Total duration of breastfeeding*						
Did not breastfeed	0	10,349	1.00 (0.98 - 1.02)	1.00 (0.98 - 1.02)	1.00 (0.98 - 1.02)	1.00 (0.98 - 1.02)
< 6 months	2.3	11,987	0.97 (0.95 - 0.98)	0.95 (0.94 - 0.97)	0.96 (0.94 - 0.98)	0.95 (0.94 - 0.97)
6 - 9 months	7.2	4,036	0.95 (0.92 - 0.98)	0.94 (0.91 - 0.97)	0.94 (0.91 - 0.97)	0.94 (0.91 - 0.97)
10 + months	18.6	5,818	0.87 (0.85 - 0.89)	0.87 (0.85 - 0.89)	0.87 (0.85 - 0.89)	0.83 (0.80 - 0.86)

* Parous women only

** Adjusted for age, socioeconomic group, BMI, smoking, alcohol, strenuous physical activity, parity, and breastfeeding as appropriate, and stratified by region

Bibliography

- [1] Stramba-Badiale M, Fox KM, Priori SG, Collins P, Daly C, Graham I, et al. Cardiovascular diseases in women: a statement from the policy conference of the European Society of Cardiology. *Eur Heart J*. 2006 Apr;27(8):994–1005.
- [2] Lewington S, Clarke R, Qizilbash N, Peto R, Collins R. Age-specific relevance of usual blood pressure to vascular mortality: a meta-analysis of individual data for one million adults in 61 prospective studies. *Lancet*. 2002 Dec;360(9349):1903–1913.
- [3] Lewington S. The importance of cholesterol, blood pressure and smoking for coronary heart disease. *European Heart Journal*. 2003 Oct;24(19):1703–1704.
- [4] Doll R, Peto R, Boreham J, Sutherland I. Mortality in relation to smoking: 50 years' observations on male British doctors. *BMJ*. 2004 Jun;328(7455):1519.
- [5] Lewington S, Clarke R. Combined Effects of Systolic Blood Pressure and Total Cholesterol on Cardiovascular Disease Risk. *Circulation*. 2005 Nov;112(22):3373–3374.
- [6] Anand SS, Islam S, Rosengren A, Franzosi MG, Steyn K, Yusufali AH, et al. Risk factors for myocardial infarction in women and men: insights from the INTERHEART study. *European Heart Journal*. 2008 Apr;29(7):932–940.
- [7] Beral V. Long term effects of childbearing on health. *J Epidemiol Community Health*. 1985 Dec;39(4):343–346.
- [8] Ness RB, Harris T, Cobb J, Flegal KM, Kelsey JL, Balanger A, et al. Number of Pregnancies and the Subsequent Risk of Cardiovascular Disease. *N Engl J Med*. 1993;328(21):1528–1533.
- [9] Green A, Beral V, Moser K. Mortality in women in relation to their childbearing history. *BMJ: British Medical Journal*. 1988 Aug;297(6645):391–395.
- [10] Palmer JR, Rosenberg L, Shapiro S. Reproductive Factors and Risk of Myocardial Infarction. *Am J Epidemiol*. 1992 Aug;136(4):408–416.
- [11] Grinblatt JA, Dekker JM, Schouten EG, Palmer JR, Rosenberg L, Shapiro S, et al. Number of Pregnancies and Risk of Cardiovascular Disease. *N Engl J Med*. 1993 Dec;329(25):1893–1895.
- [12] Parikh NI, Cnattingius S, Dickman PW, Mittleman MA, Ludvigsson JF, Ingelsson E. Parity and risk of later-life maternal cardiovascular disease. *American Heart Journal*. 2010 Feb;159(2):215–221.e6.

- [13] Schwarz EB, Ray RM, Stuebe AM, Allison MA, Ness RB, Freiberg MS, et al. Duration of lactation and risk factors for maternal cardiovascular disease. *Obstetrics and Gynecology*. 2009 May;113(5):974–982.
- [14] Stuebe AM, Michels KB, Willett WC, Manson JE, Rexrode K, Rich-Edwards JW. Duration of lactation and incidence of myocardial infarction in middle to late adulthood. *American Journal of Obstetrics and Gynecology*. 2009 Feb;200(2):138.e1–8.
- [15] ONS Office for National Statistics. Mortality statistics:General Statistics [online]; 2010. Internet [cited 02 Dec 2010]. Available from: <http://www.statistics.gov.uk/statbase/Product.asp?vlnk=618&More=Y>.
- [16] GRO General Register Office for Scotland. Scottish Mortality Statistics [online]; 2010. Internet [cited 02 Dec 2010]. Available from: <http://www.gro-scotland.gov.uk/statistics/deaths/index.html>.
- [17] NHS National Services Scotland. Information and Services Division, Scotland [online]; 2009. Internet [cited 02 Dec 2010]. Available from: http://www.isdscotland.org/isd/CCC_FirstPage.jsp.
- [18] NHS National Health Service. Health and Social Care Information Centre: Hospital Episode Statistics [online]; 2010. Internet [cited 02 Dec 2010]. Available from: <http://www.hesonline.nhs.uk>.
- [19] Kendrick S, Clarke J. The Scottish Record Linkage System. *Health Bulletin*. 1993 Mar;51(2):72–79.
- [20] World Health Organisation, editor. World Health Organization. International Statistical Classification of Diseases and Related Health Problems, vol. 1. Geneva: World Health OrganizationD; 1992.
- [21] The Million Women Study Collaborative Group. The Million Women Study: design and characteristics of the study population. *Breast Cancer Research: BCR*. 1999;1(1):73–80.
- [22] Townsend P, Phillimore P, Beattie A. Health and deprivation. Routledge; 1988.
- [23] Fletcher C, Peto R, Speizer F, Tinker C. A follow-up study of the natural history of obstructive bronchitis. Oxford University Press; 1976.
- [24] MacMahon S, Peto R, Collins R, Godwin J, MacMahon S, Cutler J, et al. Blood pressure, stroke, and coronary heart disease : Part 1, prolonged differences in blood pressure: prospective observational studies corrected for the regression dilution bias. *The Lancet*. 1990 Mar;335(8692):765–774.
- [25] Spencer EA, Appleby PN, Davey GK, Key TJ. Validity of Self-Reported Height and Weight in 4808 EPIC Oxford Participants. *Public Health Nutrition*. 2002;5(04):561–565.
- [26] Easton DF, Peto J, Babiker AGAG. Floating absolute risk: An alternative to relative risk in survival and case-control analysis avoiding an arbitrary reference group. *Statistics in Medicine*. 1991 Jul;10(7):1025–1035.
- [27] British Heart Foundation Statistics Website. Incidence of myocardial infarction [online]; 2010. Internet [cited 03 Dec 2010]. Available from: <http://www.heartstats.org/datapage.asp?id=1085>.

- [28] Fischbacher C, Bhopal R, Povey C, Steiner M, Chalmers J, Mueller G, et al. Record linked retrospective cohort study of 4.6 million people exploring ethnic variations in disease: myocardial infarction in South Asians. *BMC Public Health*. 2007;7(1):142.
- [29] Hu F. *Obesity epidemiology*. Oxford: Oxford University Press US; 2008.
- [30] Hormonal Factors in Breast Cancer CG. Alcohol, tobacco and breast cancer - collaborative reanalysis of individual data from 53 epidemiological studies, including 58 515 women with breast cancer and 95 067 women without the disease. *Br J Cancer*. 2002;87(11):1234–1245.
- [31] Yusuf S, Hawken S, Ounpuu S, Dans T, Avezum A, Lanas F, et al. Effect of potentially modifiable risk factors associated with myocardial infarction in 52 countries (the INTERHEART study): case-control study. *Lancet*. 2004 Sep;364(9438):937–952.
- [32] Kuulasmaa K, Tunstall-Pedoe H, Dobson A, Fortmann S, Sans S, Tolonen H, et al. Estimation of contribution of changes in classic risk factors to trends in coronary-event rates across the WHO MONICA Project populations. *The Lancet*. 2000 Feb;355(9205):675–687.
- [33] Promislow JHE, Gladen BC, Sandler DP. Maternal Recall of Breastfeeding Duration by Elderly Women. *Am J Epidemiol*. 2005 Feb;161(3):289–296.
- [34] Ko K, Park S, Kim Y, Bae J, Jun J, Gwack J, et al. Reliability of a questionnaire for women's reproductive history. *Journal of Preventive Medicine and Public Health Korean*. 2008;41(3).
- [35] Walshe K, Harrison N, Renshaw M. Comparison of the quality of patient data collected by hospital and departmental computer systems. *Health Trends*. 1993;25(3):105–108.
- [36] Aylin P, Bottle A, Majeed A. Use of administrative data or clinical databases as predictors of risk of death in hospital: comparison of models. *BMJ*. 2007 May;334(7602):1044.
- [37] Raftery J, Roderick P, Stevens A. Potential use of routine databases in health technology assessment. *Health Technology Assessment*. 2005 May;9(20):III–+.

Chapter 7

Childbearing and body mass index in a population of middle-aged (45 to 65 years) black South African women

Brief note on chapter content and layout

This chapter describes the design of a small cross-sectional study of the association between childbearing and BMI in a population of black South African women. The chapter begins with a review of the background and rationale for this study before describing the development of the study protocol. The chapter concludes with short summary.

7.1 Background and rationale

7.1.1 Associations between childbearing histories and BMI, and other risk factors for cardiovascular disease

Research findings presented in this thesis suggest that childbearing may influence women's BMI and through this their risk of type two diabetes. We have also seen some evidence that childbearing may influence women's risk of heart disease after the menopause. However the nature of and reproducibility of these findings in different population groups are yet to be clarified.

7.1.2 Need for more evidence from different populations of women

In exploring how my research findings in a population of UK women might be similar or different in different population groups (in particular among postmenopausal women from developing economy countries) I began by reviewing the published literature. The literature from developing economy countries on women's childbearing patterns and their non-reproductive health and risk of disease is sparse, and limited to younger women or to women living in large developed cities. [1, 2] For example, much of the literature I found from developing-economy countries is from Demographic Health and Surveillance (DHS) surveys. In these surveys reproductive history questionnaires are limited to women aged 14 to 49 years. [3? -5] In addition data on breastfeeding, if collected, is limited. [?]

7.1.3 Establishing research partnerships

As a result of my findings I began to look for ways to collaborate with researchers in developing economy countries to explore these research questions. Although I considered several developing countries in which to undertake such research I decided to work with a clinical research team lead by Professor Naomi Levitt in the Department of Medicine at the University of Cape Town, South Africa.

In 2008 Professor Naomi Levitt and Dr Krisela Steyn from the Department of Medicine at University of Cape Town, in collaboration with the Medical Research Council of South Africa established a population-based survey to update prevalence data on cardiovascular disease and associated risk factors. Within the given resource and time constraints we established that it would be feasible to design and implement a descriptive study of the reproductive histories and health profiles of women aged 45 to 65 years in parallel to the main study.

7.2 Protocol for a sub-study within CRiBSA:

Childbearing and body mass index a population of black South African women

7.2.1 The Cardiovascular Risk in Black South Africans study (CRiBSA)

The Cardiovascular Risk in Black South Africans study (CRiBSA) is a cross-sectional prevalence survey of cardiovascular risk factors in a population of black South Africans (aged between 25 and 65 years) living in the traditionally urban South African townships of the Cape Town metropolitan area. About 1 200 eligible participants have completed the survey which included interviewer-administered questionnaires which asked about cardiovascular risk factors, medical histories, socioeconomic and lifestyle factors, and other personal characteristics. Participants had their blood pressure and anthropometry measured (height, weight, waist circumference), completed an oral glucose tolerance test, and had blood collected for a biochemical assessment (glucose and lipid profiles).

A summary of the study protocol is included in Appendix C (page 252). The study was approved by the University of Cape Town Research Ethics Committee [REC REF: /224/2006]. All participants gave their written consent to take part in the study.

7.2.2 Purpose of CRiBSA sub-study

Childbearing has been shown to be associated with women's BMI and risk of BMI-associated disease (diabetes, and to a lesser extent heart disease) after the menopause

in populations of women from North America and Europe. As yet there is very little published evidence on the nature of and reproducibility of these findings in other population groups. This study will explore the association between childbearing and BMI in a population of urban-dwelling black South African women.

7.2.3 Objectives

- Describe the reproductive histories of a population of older black South African women
- Explore the association between childbearing and BMI in this population group

7.2.4 Design

Cross-sectional survey

7.2.5 Setting

The childbearing and BMI study is one of the sub-studies of the CRiBSA protocol. The parent protocol, CRiBSA, is a population-based prevalence survey of cardiovascular risk factors in black South Africans aged 25 to 65 years.

7.2.6 Participants

Women invited to enroll in the CRiBSA study and aged 45 to 65 years.

7.2.7 Exposures

Women's self-reported parity and breastfeeding histories. Women's parity will be summarised into the categories nulliparous, or for parous women quartiles of parity using the mean and standard deviation of the survey data. Total duration of breastfeeding in months will be modeled as a continuous and categorical variable. The distribution of the data (mean and standard deviation) will be used to calculate quartiles of total breastfeeding duration in months (excluding parous women who did not breastfeed).

7.2.8 Main outcome measure

Mean measured BMI (kg/m^2).

7.2.9 Methods

7.2.9.1 Data Analysis

Analyses will be conducted using standard statistical methods. Data will be described using the mean and standard deviation for continuous variables and frequencies for categorical variables. Linear regression models will be used to explore the association between women's childbearing histories and their mean BMI. All models will be adjusted for known and potential co-factors.

7.2.9.2 Sample size

Based on results from Million Women Study analyses of the effects of parity and breastfeeding on women's BMI after the menopause I calculated the sample sizes required

to detect a difference in BMI between the different categories of parity and of breast-feeding of between 0.5 kg/m² and 1.5 kg/m². The mean BMI of black South African women aged 45 to 65 years, living in Cape Town, is estimated to be about 28.1 kg/m² (SD 2.9). Table 7.1(page 222) shows the estimated sample sizes required to detect 0.5 to 1.5 kg/m² difference in BMI (with a standard deviation of 1.5 kg/m² or 3 kg/m²) between two groups with 90% power with an alpha <0.05. Considering the constraints of adding a parallel study on to an existing study, and based on findings of the change in BMI associated with parity and breastfeeding in a large population of UK women (see chapter 4, page 87) it is reasonable to power the study to assess the association between look for a minimum change in mean BMI of 0.7 kg/m² associated with changes in the exposure groups.

To determine the relative size of the groups I used data from hospital-based controls of a lactation and breast cancer study carried out in Cape Town in the 1990s to estimate the expected childbearing patterns. [6] I was unable to find any published population-based data on black South African women's childbearing histories. Table 7.2 (page 222) shows selected childbearing characteristics for women in the control arm of the study. From the table it can be seen that roughly 10% of women in the control arm were nulliparous, nearly one quarter had a parity of three, and 40% had ≥ 4 children. 15% of parous women did not breastfeed.

To have sufficient power to detect a 0.7 kg/m² difference in BMI between two parity groups would required 97 women per group (if the standard deviation for BMI was 1.5 kg/m²) or 386 per group (if the standard deviation for BMI was 3 kg/m²). Given the parity distribution I would aim to recruit about 100 nulliparous women, the total number of women recruited (roughly 1000) would provide sufficient power to power to detect a 1 kg/m² change in BMI between different quartiles of parity and breastfeeding

duration if the BMI variability is high, and up to a 0.5 kg/m² change in BMI between two quartiles if BMI variability is lower.

7.2.10 Questionnaire design, training materials, data collection and monitoring

7.2.10.1 Co-ordination and collaboration

This study will be conducted in collaboration with Professor Naomi Levitt and Dr Krisela Steyn, the study co-ordinating centre will be based in Department of Endocrinology at the University of Cape Town (Old Main Building, Groote Schuur).

7.2.10.2 Questionnaire design

In association with senior scientists from the Million Women Study and from the University of Cape Town I designed a locally relevant reproductive history questionnaire. I began by reviewing the design and development of the questionnaires used in the Million Women Study and the interviewer administered questionnaires already designed for the CRiBSA study. I was mindful that although this questionnaire would be administered by a trained interviewer, it was going to be used in a busy research clinic as part of a package of at least four questionnaires one of which was at least 20 pages long. This questionnaire needed to be easy to use, quick to fill in, easy to review for data inconsistencies, and it needed to comprehensively capture women's reproductive histories. I based the format of the questionnaire on that of the Million Women Study questionnaires limiting the length of the questionnaire to two double-sided A4 pages. The questionnaire was divided into four interviewer completed sections and two review sections to be completed by the clinic coordinator. The questionnaire was

designed in English. I worked with Sister Buyelwa Majikela-Dlangamandla and Sister Theresa Goga to translate and back translate the questionnaire from English into isiXhosa making changes to the wording of the questions as appropriate.

7.2.10.3 Questionnaire pilot

I piloted the English and isiXhosa versions of the questionnaire using female volunteers from a pensioners lunch club in Nyanga. I reviewed the results of the pilot with Sister Buyelwa Majikela-Dlangamandla and Sister Theresa Goga and we worked together to develop a final version of the questionnaire. Copies of the questionnaire in English can be found in Appendix D (page 255).

7.2.10.4 Training materials

I have developed written training and reference materials on the questionnaires for interviewers to use. With assistance from Sister Theresa Goga and Sister Debbie Jonathan I will provide the training in the use of the questionnaire to the interviewers. Senior clinic staff including the clinic coordinator will also be given training in how to use the training materials, how to administer the questionnaires, and how to double check the information recorded on the questionnaire. A copy of the training materials can also be found in Appendix D (page 264).

7.2.10.5 Data collection

Trained interviewers will administer the reproductive history questionnaire in conjunction with the three other previously validated questionnaires used in the main study.

At the clinic visit data are collected on socio-demographic variables, dietary intake, exercise, alcohol consumption, tobacco use, and psychological stress. [7–11] Biological data including anthropometry and blood are also collected.

At every clinic both English and isiXhosa versions of the questionnaire will be available, as will a reference training manual, and at least one senior member of the field team who can answer questions or queries interviewers or participants might have.

7.2.10.6 Data monitoring and quality checks

At the end of the clinic, before participants leave, the questionnaires will be reviewed independently by another team member for completion and accuracy of information (for example ensuring correct chronology of events). Any problems identified will be resolved through discussion with the interviewer and if necessary with a repeat interview of the participant.

To assess how well women are able to recall their reproductive histories a 10% sample of randomly selected women will be asked if they can be contacted for a repeat interview. The same questionnaire will be used for the repeat interview. The interval between questionnaires will be one month or less. Participants will be contacted via telephone and the repeat interview will be conducted in-person.

7.2.10.7 Data capture and storage

Each questionnaire will be coded using a standard format. Data from coded questionnaires will be anonymised and double entered into a secure study database. All of the electronic data will be stored securely on an encrypted computer. Hard copies of the

completed reproductive history questionnaires will be kept in locked offices at the University of Cape Town. The health profile and biological data will be stored separately at the Medical Research Council offices in Cape Town. The health profile and biological data are also anonymised and entered into a separate secure database. A copy of this data will be sent securely to the Cancer Epidemiology Unit (Oxford) and merged with data from the reproductive history questionnaires. The merged dataset will be checked for logical and other inconsistencies. This data will be stored securely on an encrypted computer access to which is restricted to named researchers.

7.3 Practical procedures

7.3.1 Recruitment

All women between the ages of 45 and 65 years living in households selected for the CRiBSA study will be invited participate in the reproductive histories study. Women who agree to participate will be invited to attend one of the CRiBSA research clinics. At the clinic the study and the questionnaire will again be explained in the participants first language. Women will be asked for written consent to participate in the study. Women who agree to participate will have a reproductive history questionnaire added to their research clinic folder (which will contain all of the CRiBSA questionnaires and a clinic checklist) and a green sticker placed on top (to facilitate identification and subsequent data at the main research office).

7.3.2 Questionnaire administration

In order to minimise disruption to the flow of participants through the research clinic the reproductive histories questionnaire will most likely be administered during the interval between the second and third blood draw (as part of the oral glucose tolerance test protocol) and before the dietary recall questionnaire is administered.

7.4 Discussion

Summary

The planned study is an extension of work presented in earlier chapters in this thesis. The purpose of the chapter is to begin clarifying how childbearing might influence women's mean BMI after their reproductive years are completed in a different population group. As part of this I described a study protocol examining the association between childbearing and BMI in a population of middle-aged black South African women. I also described the design and pilot of a locally relevant reproductive history questionnaire. Using the findings from this thesis and the available published data I calculated several sample size estimates to adequately power such a study. I have also considered how the expected patterns of childbearing in this population group will affect the number of women I would need to recruit to achieve this.

7.5 Chapter summary

This chapter is informed by research findings presented in earlier chapters, it extends this work by exploring how the associations between women's childbearing and BMI

(and through this their risk of diabetes mellitus and ischaemic heart disease) after the menopause may differ across different population groups. The chapter describes preliminary thinking and work carried out in order to develop an appropriate study protocol to address this research question. Data required to present analysis and results for this chapter are awaited.

TABLE 7.1: Sample sizes required to adequately power a study designed to detect a change in the mean BMI between two exposure groups

Change in mean BMI (kg/m²) (mean BMI 28.1 [SD 3])	Category A	Category B
0.5 kg/m²	757	757
0.7 kg/m²	386	386
1 kg/m²	190	190
1.5 kg/m²	85	85
(mean BMI 28.1 [SD 1.5])	Category A	Category B
0.5 kg/m²	190	190
0.7 kg/m²	97	97
1 kg/m²	48	48
1.5 kg/m²	22	22

TABLE 7.2: Selected characteristics of women participating in the control arm of a Cape Town-based case-control study of the association between lactation and risk of breast cancer (Coogan et al 1999)

Population Characteristics	All Women
	n = 1625
Age ≥ 40 years [n] (%)	845 (57)
Parity [n] (%)	
0	154 (9)
1	234 (15)
2	316 (21)
3	334 (23)
≥4	587 (41)
Women who breastfed [n] (%)	[1250] (85)
Average number of children breastfed	3
Duration breastfeeding	
Total duration in months (Mean)	20.0
Duration per child in months (Mean)	11.9

* Amongst parous women only

Bibliography

- [1] Coitinho DC, Sichieri R, Benício MHD. Obesity and Weight Change Related to Parity and Breast-Feeding Among Parous Women in Brazil. *Public Health Nutrition*. 2001;4(04):865–870.
- [2] Villegas R, Gao Y, Yang G, Li HL, Elasy T, Zheng W, et al. Duration of breast-feeding and the incidence of type 2 diabetes mellitus in the Shanghai Women’s Health Study. *Diabetologia*. 2008 Feb;51(2):258–266.
- [3] ICF Macro: an ICF International Company . Demographic & Health Research [online]; 2011. Internet [cited 02 Feb 2011]. Available from: <http://www.macrominternational.com/programareas/Health/research.aspx>.
- [4] Kim SA, Yount KM, Ramakrishnan U, Martorell R. The relationship between parity and overweight varies with household wealth and national development. *International Journal of Epidemiology*. 2007 Feb;36(1):93–101.
- [5] Kim SA, Stein AD, Martorell R. Country development and the association between parity and overweight. *Int J Obes*. 2006 Oct;31(5):805–812.
- [6] Coogan PF, Rosenberg L, Shapiro S, Hoffman M. Lactation and breast carcinoma risk in a South African population. *Cancer*. 1999 Sep;86(6):982–989.
- [7] Ewing JA. Detecting Alcoholism. *JAMA: The Journal of the American Medical Association*. 1984 Oct;252(14):1905–1907.
- [8] Levitt NS, Katzenellenbogen JM, Bradshaw D, Hoffman MN, Bonnici F. The prevalence and identification of risk factors for NIDDM in urban Africans in Cape Town, South Africa. *Diabetes Care*. 1993 Apr;16(4):601–607.
- [9] Vorster HH, Jerling JC, Steyn K, Badenhorst CJ, Slazus W, Venter CS, et al. Plasma fibrinogen of black South Africans: the BRISK study. *Public Health Nutrition*. 1998 Sep;1(03):169–176.
- [10] Lambert EV, Steyn K, Stender S, Everage N, Fourie JM, Hill M. Cross-cultural validation of the hill-bone compliance to high blood pressure therapy scale in a South African, primary healthcare setting. *Ethnicity & Disease*. 2006;16(1):286–291.
- [11] Dennison CR, Peer N, Lombard CJ, Kepe L, Levitt NS, Steyn K, et al. Cardiovascular risk and comorbid conditions among Black South Africans with hypertension in public and private primary care settings: the HiHi study. *Ethnicity & Disease*. 2007;17(3):477–483.

Chapter 8

Discussion

8.1 Summary of main findings

Background

Worldwide the greatest number and proportion of older people are women. [1] This population group presents unique challenges to current social and health care systems. Older populations suffer a disproportionate burden of chronic disease and life lived with a disability contributing to the growing cost burden of health care. [1] Research into health risks and modifiable disease risk factors at a population level is a pre-requisite to galvanising appropriate health resources for this population group.

For common conditions including certain cancers childbearing (specifically parity and breastfeeding) has been shown to have a persistent effect on women's risk of disease. [2–10]. For other common conditions affecting older women, excess adiposity, type II diabetes mellitus, and heart disease the influence of common exposures like childbearing on women's long-term risks of disease has been unclear.

The literature I reviewed and discussed in Chapters one and two highlighted the continued uncertainty of the associations between childbearing and these common conditions and important limitations in current published evidence base; lack of exposure data (especially for breastfeeding) and therefore the small size of the studies for which data are available, and the limited ability to control for important co-factors and important potential effect mediators. Given the lack of clear evidence of the effects of a very common exposure on increasingly common outcomes, particularly where this exposure has already been found to affect other common outcomes (for example, gall

bladder related diseases), and the potential public health impact in terms of number of women affected the primary aim of my thesis was to explore the hypothesis that childbearing (specifically parity and breastfeeding) is associated with women's body weight and risk of excess adiposity, and also with women's risk of diabetes mellitus (type II), and ischaemic heart disease in the long-term (i.e. even after her reproductive years are complete).

Main findings from this thesis

Data quality

I established the quality and completeness of the data available for analysis from the Million Women Study. Most women who were asked about their childbearing histories reported on them (see Appendix B, page 250) I did not find evidence that women who did not report this information were substantially different. Outcome data used in this thesis came from three sources. For BMI I used self-reported height and weight data, in sensitivity analyses I found that BMI from self-reported data compared well to BMI calculated from measured height and weight data. For diabetes I used self-reported diagnosis of diabetes at first repeat survey among women who had not previously reported a diagnosis of diabetes. Although there is the potential for misclassification work done previously by my colleague Dr Richard Stevens has shown that women who have had a diagnosis of diabetes will tend to report it (sensitivity is high). For ischaemic heart disease I used outcome data from hospital admissions records and death certificates. Data captured from these two sources has been found to be of good quality. In addition I found that the age-standardised incidence rates for IHD for Million Women Study participants were comparable to other population-based data from the

UK. Information on various covariates I considered and explored were, for time unchanging variables, taken from the recruitment questionnaire and for time changing covariates I used the regression dilution technique to account for changes in women's exposure status over time.

Breastfeeding trends in the UK over time

Breastfeeding data from the Million Women Study are highly comparable to data from the 1958 and 1970 birth cohorts, and the early, National Infant Nutrition Feeding surveys. For each time point examined the proportions of mothers from Million Women Study who reported breastfeeding were remarkably similar to women from the other studies. For example, 36% of Million Study women with a child born in 1970 reported breastfeeding compared to 37% for women participating in the 1970 birth cohort.

All of the data suggest that breastfeeding trends in the United Kingdom have varied considerably during the last half of the twentieth century. Specifically, the proportion of women breastfeeding declined from around two thirds of parous women in the 1950s to only about one third of parous women in the early 1970s. Thereafter the trend was reversed and by the mid 1980s the proportion of women who reported breastfeeding was similar (though still slightly lower) to that of the 1950s. Differences by socioeconomic group and birth order were noted but did not affect the overall pattern. There is some evidence to suggest that while the overall proportion of women who breastfeed has increased since the 1970s the difference in the proportion of women who breastfeed between the highest and lowest socioeconomic groups has increased.

Effects of childbearing on women's body weight

Persistent and independent associations between parity, and breastfeeding history and mean BMI were found in an analysis which included almost 750,000 postmenopausal women in the UK. The standardised mean BMI increased progressively with number of births from 25.6 kg/m² in nulliparous women up to 27.2 kg/m² for women with four or more births, a difference of 1.7 (95% confidence interval 1.6 to 1.7) kg/m². At every parity level the standardised mean BMI was lower among women who had breastfed. The mean BMI decreased the longer women had breastfed and was 0.5 (0.5 to 0.6) kg/m² lower in those who had breastfed for a total of 10 months or more (average 18 months) than in women who had never breastfed. These differences were statistically significant ($p < 0.0001$) and independent of the effects of socioeconomic group, region, smoking, and physical activity. Although the effect size appeared modest it was noted that the associations between parity, breastfeeding and BMI in postmenopausal women of a similar order of magnitude to established risk factors known to be associated with BMI, for example smoking, and physical activity.

Effects of childbearing on women's risk of diabetes mellitus (type II)

In the analysis of half-a-million women and almost 8000 cases of diabetes mellitus I found evidence that the risk of diabetes mellitus after the menopause associated with childbearing is largely mediated through BMI. However adjusting for BMI did not fully explain the associations found with increased parity (women of parity $\geq$ three had 20% increased risk - OR 1.20, CI 1.08 to 1.34) and having ever breastfed (11% lower risk - OR 0.89, CI 0.84 to 0.94). It is worthwhile noting that in this study population I did not find evidence for a trend with longer durations of breastfeeding ($p = 0.9654$) which may in

part be due to the short duration of total breastfeeding for most of the women in this study (and therefore limited power to assess the effects of long total durations.)

Effects of childbearing on women's risk of IHD

In this analysis which included 35,590 incident events during 7,735,000 person-years of follow-up time increasing parity (≥ 3) was statistically and meaningfully associated with a higher risk of IHD. After adjusting for known important risk factors associated with IHD such as age and smoking nulliparous women had the lowest risk of incident IHD (RR 0.90, CI 0.87 to 0.93). It is known that nulliparous and parous women are quite different especially in a generation where reproductive choice was more limited than it is at present. It is also known that childbearing in itself may be a marker for health. As such the finding that nulliparous women have a different risk of incident IHD may not be surprising however I think it adds weight to the argument that in studies exploring the associations between childbearing and risk of disease nulliparous women should not be used as the reference group.

Among parous women, parity three or more was associated with a modest increase in the risk of incident IHD (12% excess risk at parity level three (RR 1.12, CI 1.09 to 1.16) and 31% at parity four or more (RR 1.31, CI 1.27 to 1.36)). For parous women a history of breastfeeding was associated with a small decrease in the risk (RR 0.92, CI 0.90 to 0.95) and there was a small but statistically significant trend associated with duration of breastfeeding (RR 0.96 per additional month of breastfeeding, CI 0.95 to 0.98). Although the associations between childbearing and IHD are small in comparison to known important risk factors for example smoking or BMI the effects are statistically significant and clinically meaningful.

Findings in context

The findings presented in this thesis add new information to the literature on the direction of associations between childbearing (specifically parity and breastfeeding) and women's adiposity (as captured by BMI), and risk of diabetes mellitus (type II), or incident ischaemic heart disease after the menopause and the likely causal nature of the associations I found. They clarify the independent and joint effects of parity and breastfeeding on the outcomes of interest and they put in to context the relative importance of these exposures compared to known key risk factors, thereby addressing the original aim and the specific objectives of this thesis.

To my knowledge the work in chapter 3 is the first time that historical trends in breastfeeding have been explored in the UK, and that the first time infant nutrition data from several independent sources have been used to explore the quality of long-term maternal recall of breastfeeding. It is hoped that this knowledge will contribute to the collection and use of such data in future epidemiological studies.

Strengths and weaknesses

Strengths of this study include its size and the availability of information about co-factors. The Million Women Study includes one in four UK women aged between 50 and 64 years at the time of recruitment. Detailed data are available for childbearing and breastfeeding, and for key co-factors including socioeconomic factors, smoking and physical activity. Updated information on important co-factors are available from the first repeat survey (about three years after recruitment). Outcomes assessed in this

thesis were obtained from various sources including self-report, central registries (Office of National Statistics), and hospital admissions databases from Scotland (Scottish Morbidity Record) and from England (Hospital Episode Statistics, HES).

As with most large-scale epidemiological studies exposure information was obtained by self-report. Studies in older women have found that their long-term recall of parity and breastfeeding is reliable, [11, 12] and I have shown previously that self-reported breastfeeding data corresponds well to data obtained from historical data sets (Chapter 3, page 63). For the comparison of breastfeeding data from different sources I was limited by the nature of the different sources and types of data. For example, the birth cohorts and feeding surveys were designed to answer questions about infants rather than mothers. I was also limited in my ability to compare duration of breastfeeding across studies because of the different ways in which duration had been assessed.

For BMI (kg/m^2) as an outcome I used self-reported height and weight data to calculate BMI (kg/m^2). Such data may be affected by random and systematic measurement error, however previous studies within similar population groups have shown that this is unlikely to be a source of major material bias. [13, 14] I found a high degree of correlation when I compared BMI calculated from measured weight and height data with self-reported data, and in a sensitivity analysis there was no meaningful difference in the results obtained using measured BMI data compared to self-reported data. Where BMI (kg/m^2) was used as a co-factor I used the regression dilution approach to correct for possible misclassification. [15]

For the outcome diabetes mellitus (type II) I used self-reported diagnosis of or treatment for diabetes mellitus. In chapter 2 (page 18) I showed evidence for the reliability and validity of self-reported diabetes in this study population, these findings agree

with the published literature. [16–23]

For the analysis of risk of ischaemic heart disease I used outcome data from hospital episode statistics and death certificates. In the UK death certification data are reliable and complete. [24] Data from hospital admissions has been shown to be reasonably reliable. [25–27] As the hospital admissions databases started in from January 1981 in Scotland and April 1997 in England it is possible that a woman's first admission with IHD during the Million Women Study follow-up period was not her first IHD event. [28, 29] If a true effect of childbearing existed on a woman's risk of a first ischaemic event it could be masked by the increase in risk of a second IHD event (56,000 women in this study population reported a diagnosis of or treatment for heart disease at recruitment) again though any effect would likely still be small and overshadowed by known risk factors.

8.2 Possible implications

An association between between childbearing and subsequent risk of disease has now be shown for several conditions including cancers (primarily of reproductive organs), symptomatic gall bladder disease, BMI, and possibly IHD. [2, 3, 6, 10] In addition in this thesis I have shown a strong dose-dependent association between childbearing and BMI that in large part mediates the risks associated between childbearing and subsequent risk of diabetes. I have also shown a modest dose-dependent association between childbearing and women's risk of IHD that is independent of the risk of IHD associated with BMI and additional co-factors. However, it is known that the relationship between the exposures (parity and breastfeeding) and long-term non-reproductive outcomes is complicated and may be mediated by various factors which may in part

be behavioural (although there is a growing body of basic-science research demonstrating long-term effects of childbearing associated hormones like oxytocin and prolactin on the cardiovascular system.) It remains prudent to wait until the findings presented in this thesis have been replicated in other population groups, and in other conditions with a similar pathophysiology (for example comparing the results for IHD to those found with ischaemic stroke) before discussing potential public health implications in depth.

8.2.1 Younger women

Good evidence already exists for the association between reduced family size, long durations of breastfeeding and improved health and other outcomes for mothers and their children (the WHO recommends about 6 months of exclusive breast feeding or about 12 months breastfeeding in total). [30, 31] [32] Given the conflicting published evidence on the short-term effects of breastfeeding on women's weight postpartum, and the low risk of IHD in this population it is prudent to interpret our results cautiously, and to base policy recommendations on the existing evidence of paediatric and maternal benefits (or at least lack of harm) of breastfeeding.

8.2.2 Older women

Although findings from this thesis clearly showed an association between childbearing and adiposity, risk of diabetes (again mediated through the effects on adiposity), and risk of IHD; the results showed that contemporary risk factors (for example smoking, or physical activity) are at least (in the case of adiposity) if not more (in the case of diabetes and IHD) important. Certainly for women who are older it is important to

advocate for modifying these key risk factors (many of which are shared risk factors for several conditions).

8.3 Future research

8.3.1 Mechanism of action

Further basic science research is required to explore possible biological mechanisms underpinning these associations including the role of centrally acting hormones (prolactin and oxytocin; at this stage oxytocin is the more promising of the two as it has been found to be more closely correlated to suckling and lactation than prolactin, and it has been found in animal and human studies to lower blood pressure among other effects) [33–36], peripherally acting hormones (cortisol), [37] and tissue specific effects (for example, different stores of adipose tissue). [38]

8.3.2 Other outcomes

Childbearing has been suggested to influence the risk of other maternal outcomes for example stroke. [39–42] The Million Women Study contains sufficient data for additional analyses exploring the associations between childbearing and stroke, as well as other cardiovascular outcomes.

8.3.3 Generalisability of findings

Data are now available, and preliminary findings have been published from several large cohorts of women in various countries including Australia, the USA, China, Korea, as well as several European countries. [43–47] A meta-analysis of all the available data on the associations between childbearing and various non-cancer outcomes in older women would be extremely useful. Data from women who live in resource poor settings and who in general have much higher parities and longer durations of breastfeeding are still awaited (the research studies in China and Korea are based in large well-resourced urban settings). Data from the cross-sectional survey in a population of black South African women (predominately living in poor urban settings) are awaited for analysis and will hopefully add to the existing literature.

8.4 Conclusions

In a large population of UK women childbearing was found to have a persistent influence on women's adiposity after the menopause, and through this the risk of diabetes mellitus. Childbearing was also found to be modestly associated with women's risk of IHD. The associations between childbearing and BMI were of a similar order of magnitude to risk factors known to be associated with BMI. In contrast, any independent associations between childbearing and risk of diabetes or risk of IHD after the menopause were found to be small in comparison to known important risk factors. Future research will clarify the extent to which childbearing is associated with risk for other non-reproductive outcomes (for example stroke), explore for potential biological mechanisms of action, and establish the extent to which findings presented in this thesis are applicable to other population groups.

Bibliography

- [1] Goulding M, Rogers M, Smith S. Public Health and Aging: Trends in Aging — United States and Worldwide [online]; 2010. Internet [cited 11 Nov 2010]. Available from: <http://www.cdc.gov/mmwr/preview/mmwrhtml/mm5206a2.htm>.
- [2] Beral V. Long term effects of childbearing on health. *J Epidemiol Community Health*. 1985 Dec;39(4):343–346.
- [3] Adami H, Lambe M, Persson I, Ekblom A, Adami HO, Hsieh CC, et al. Parity, age at first childbirth, and risk of ovarian cancer. *The Lancet*. 1994 Nov;344(8932):1250–1254.
- [4] Green A, Beral V, Moser K. Mortality in women in relation to their childbearing history. *BMJ: British Medical Journal*. 1988 Aug;297(6645):391–395.
- [5] Doblhammer G. Reproductive history and mortality later in life: A comparative study of England and Wales and Austria. *Population Studies: A Journal of Demography*. 2000 Jul;54(2):169–176.
- [6] Collaborative Group on Hormonal Factors in Breast Cancer. Breast cancer and breast-feeding: collaborative reanalysis of individual data from 47 epidemiological studies in 30 countries, including 50 302 women with breast cancer and 96 973 women without the disease*. *The Lancet*. 2002 Jul;360(9328):187–195.
- [7] Beral V, Banks E, Reeves G. Evidence from randomised trials on the long-term effects of hormone replacement therapy. *Lancet*. 2002 Sep;360(9337):942–944.
- [8] Doblhammer G, Oeppen J. Reproduction and longevity among the British peerage: the effect of frailty and health selection. *Proceedings of the Royal Society B: Biological Sciences*. 2003 Aug;270(1524):1541–1547.
- [9] Grundy E. Women's fertility and mortality in late mid life: A comparison of three contemporary populations. *American Journal of Human Biology: The Official Journal of the Human Biology Council*. 2009 Aug;21(4):541–547.
- [10] Liu B, Beral V, Balkwill A. Childbearing, breastfeeding, other reproductive factors and the subsequent risk of hospitalization for gallbladder disease. *International Journal of Epidemiology*. 2009 Feb;38(1):312–318.
- [11] Promislow JHE, Gladen BC, Sandler DP. Maternal Recall of Breastfeeding Duration by Elderly Women. *Am J Epidemiol*. 2005 Feb;161(3):289–296.
- [12] Ko K, Park S, Kim Y, Bae J, Jun J, Gwack J, et al. Reliability of a questionnaire for women's reproductive history. *Journal of Preventive Medicine and Public Health Korean*. 2008;41(3).

- [13] Spencer EA, Roddam AW, Key TJ. Accuracy of self-reported waist and hip measurements in 4492 EPIC-Oxford participants. *Public Health Nutrition*. 2004 Sep;7(6):723–727.
- [14] Whitlock G, Clark T, Hoorn SV, Rodgers A, Jackson R, Norton R, et al. Random errors in the measurement of 10 cardiovascular risk factors. *European Journal of Epidemiology*. 2001;17(10):907–909.
- [15] Fletcher C, Peto R, Speizer F, Tinker C. A follow-up study of the natural history of obstructive bronchitis. Oxford University Press; 1976.
- [16] Banks E, Beral V, Cameron R, Hogg A, Langley N, Barnes I, et al. Agreement between general practice prescription data and self-reported use of hormone replacement therapy and treatment for various illnesses. *Journal of Epidemiology and Biostatistics*. 2001;6(4):357.
- [17] Heliövaara M, Aromaa A, Klaukka T, Knekt P, Joukamaa M, Impivaara O. Reliability and validity of interview data on chronic diseases The mini-Finland health survey. *Journal of Clinical Epidemiology*. 1993 Feb;46(2):181–191.
- [18] Norrish A, North D, Kirkman P, Jackson R. Validity of Self-reported Hospital Admission in a Prospective Study. *American Journal of Epidemiology*. 1994 Nov;140(10):938 –942.
- [19] Bowlin SJ, Morrill BD, Nafziger AN, Lewis C, Pearson TA. Reliability and changes in validity of self-reported cardiovascular disease risk factors using dual response: The behavioral risk factor survey. *Journal of Clinical Epidemiology*. 1996 May;49(5):511–517.
- [20] The Italian Longitudinal Study on Aging Working Group. Prevalence of chronic diseases in older Italians: comparing self-reported and clinical diagnoses. *International Journal of Epidemiology*. 1997 Oct;26(5):995 –1002.
- [21] Goldman N, Lin I, Weinstein M, Lin Y. Evaluating the quality of self-reports of hypertension and diabetes. *Journal of Clinical Epidemiology*. 2003 Feb;56(2):148–154.
- [22] Okura Y, Urban LH, Mahoney DW, Jacobsen SJ, Rodeheffer RJ. Agreement between self-report questionnaires and medical record data was substantial for diabetes, hypertension, myocardial infarction and stroke but not for heart failure. *Journal of Clinical Epidemiology*. 2004 Oct;57(10):1096–1103.
- [23] Simpson CF, Boyd CM, Carlson MC, Griswold ME, Guralnik JM, Fried LP. Agreement Between Self-Report of Disease Diagnoses and Medical Record Validation in Disabled Older Women: Factors That Modify Agreement. *Journal of the American Geriatrics Society*. 2004;52(1):123–127.
- [24] ONS Office for National Statistics. Mortality statistics:General Statistics [online]; 2010. Internet [cited 02 Dec 2010]. Available from: <http://www.statistics.gov.uk/statbase/Product.asp?vlnk=618&More=Y>.
- [25] Walshe K, Harrison N, Renshaw M. Comparison of the quality of patient data collected by hospital and departmental computer systems. *Health Trends*. 1993;25(3):105–108.
- [26] Aylin P, Bottle A, Majeed A. Use of administrative data or clinical databases as predictors of risk of death in hospital: comparison of models. *BMJ*. 2007 May;334(7602):1044.

- [27] Raftery J, Roderick P, Stevens A. Potential use of routine databases in health technology assessment. *Health Technology Assessment*. 2005 May;9(20):III–+.
- [28] Kendrick S, Clarke J. The Scottish Record Linkage System. *Health Bulletin*. 1993 Mar;51(2):72–79.
- [29] NHS National Health Service. Health and Social Care Information Centre: Hospital Episode Statistics [online]; 2010. Internet [cited 02 Dec 2010]. Available from: <http://www.hesonline.nhs.uk>.
- [30] Kramer MS, Kakuma R. Optimal duration of exclusive breastfeeding. *Cochrane Database of Systematic Reviews (Online)*. 2002;(1):CD003517.
- [31] Ip S, Chung M, Raman G, Chew P, Magula N, DeVine D, et al. Breastfeeding and maternal and infant health outcomes in developed countries. *Evidence Report/Technology Assessment*. 2007 Apr;(153):1–186.
- [32] World Health Organisation. WHO - Breastfeeding; 2011. Internet [cited 02 Feb 2011]. Available from: <http://www.who.int/topics/breastfeeding/en/>.
- [33] Heinig MJ, Dewey KG. Health Effects of Breast Feeding for Mothers: A Critical Review. *Nutrition Research Reviews*. 1997;10(01):35–56.
- [34] Peterson M, Alster P, Lundeburg T, Uvnäs-Moberg K. Oxytocin Causes a Long-Term Decrease of Blood Pressure in Female and Male Rats. *Physiology & Behavior*. 1997 Nov;60(5):1311–1315.
- [35] Hill PD, Chatterton RT, Aldag JC. Serum Prolactin in Breastfeeding: State of the Science. *Biological Research For Nursing*. 1999 Jul;1(1):65–75.
- [36] Petersson M, Hulting A, Uvnäs-Moberg K. Oxytocin causes a sustained decrease in plasma levels of corticosterone in rats. *Neuroscience Letters*. 1999 Apr;264(1-3):41–44.
- [37] Lankarani-Fard A, Kritz-Silverstein D, Barrett-Connor E, Goodman-Gruen D. Cumulative Duration of Breast-Feeding Influences Cortisol Levels in Postmenopausal Women. *Journal of Women's Health & Gender-Based Medicine*. 2004 Jul;10(7):681–687.
- [38] Lassek WD, Gaulin SJ. Changes in body fat distribution in relation to parity in American women: a covert form of maternal depletion. *American Journal of Physical Anthropology*. 2006 Oct;131(2):295–302.
- [39] Zhang X, Shu X, Gao Y, Yang G, Li H, Zheng W. Pregnancy, Childrearing, and Risk of Stroke in Chinese Women. *Stroke; a Journal of Cerebral Circulation*. 2009 May;.
- [40] Mhurchu CN, Anderson C, Jamrozik K, Hankey G, Dunbabin D. Hormonal factors and risk of aneurysmal subarachnoid hemorrhage: an international population-based, case-control study. *Stroke; a Journal of Cerebral Circulation*. 2001 Mar;32(3):606–612.
- [41] Jung S, Bae H, Park B, Yoon B. Parity and risk of hemorrhagic strokes. *Neurology*. 2010 May;74(18):1424–1429.

- [42] Yang L, Kuper H, Sandin S, Margolis KL, Chen Z, Adami H, et al. Reproductive history, oral contraceptive use, and the risk of ischemic and hemorrhagic stroke in a cohort study of middle-aged Swedish women. *Stroke; a Journal of Cerebral Circulation*. 2009 Apr;40(4):1050–1058.
- [43] Yoo KY, Shin HR, Chang SH, Lee KS, Park SK, Kang D, et al. Korean Multi-center Cancer Cohort Study including a Biological Materials Bank (KMCC-I). *Asian Pacific Journal of Cancer Prevention: APJCP*. 2002;3(1):85–92.
- [44] Zheng W, Chow W, Yang G, Jin F, Rothman N, Blair A, et al. The Shanghai Women's Health Study: rationale, study design, and baseline characteristics. *American Journal of Epidemiology*. 2005 Dec;162(11):1123–1131.
- [45] Schwarz EB, Ray RM, Stuebe AM, Allison MA, Ness RB, Freiberg MS, et al. Duration of lactation and risk factors for maternal cardiovascular disease. *Obstetrics and Gynecology*. 2009 May;113(5):974–982.
- [46] Stuebe AM, Michels KB, Willett WC, Manson JE, Rexrode K, Rich-Edwards JW. Duration of lactation and incidence of myocardial infarction in middle to late adulthood. *American Journal of Obstetrics and Gynecology*. 2009 Feb;200(2):138.e1–8.
- [47] Liu B, Jorm L, Banks E. Parity, breastfeeding, and the subsequent risk of maternal type 2 diabetes. *Diabetes Care*. 2010 Jun;33(6):1239–1241.

Appendix A

Million Women Study questionnaires

Sample copies of the recruitment questionnaires (known as the “blue” questionnaire discussed in Chapter 2, 19) and the first resurvey questionnaire (known as the “yellow” questionnaire).

A.1 Recruitment questionnaire (blue)

THE MILLION WOMEN STUDY

A national survey of women invited for breast screening

We need one million women to help us in research that will benefit women all over the world.

Would you become one of these special women?

More and more women are taking hormone replacement therapy (HRT) so it is vital that we find out as much as possible about its benefits and any possible side effects. We have a unique opportunity through the NHS Breast Screening Programme to learn about the way different types of HRT and other lifestyle factors affect a woman's health, particularly her breasts. Britain is the only country in the world that can carry out this study because it is the only one with the combination of a large population and a comprehensive national breast screening programme.

The NHS Breast Screening Programme, the Imperial Cancer Research Fund and the Medical Research Council have joined together to organise The Million Women Study. If one million women answer this questionnaire over the next three years we could have some of the answers to our most important questions about HRT within five years or so.

We would be very grateful if you could set aside some time to answer these questions. It should not take more than 10-15 minutes. You do not have to answer this questionnaire and if you decide not to you will still have your screening done in the normal way.

Please answer every question and do not leave blanks as all the information that you give us is very useful. If you are not sure about exact dates or ages an approximate answer is better than none. If you have any questions you can ring us on freephone 0800 262 872.

Even if you are not taking HRT it is just as important that you fill in the questionnaire.

Please bring this questionnaire to your breast screening appointment

To help us read your answers please write as clearly as possible and be sure to complete the questionnaire as shown:

Please put a cross in the appropriate box(es) ☒ ☐ ☐

OR put numbers in the appropriate box e.g. 23rd April 1946 age years

GENERAL QUESTIONS ABOUT YOU

1. What is your date of birth? (please put day/month/year)

 / /

2. How old are you? years

3. How tall are you? (please give to the nearest inch)

 feet inches

4. About how much do you weigh?

 stone lbs

5. How old were you when you finished full time schooling? (please cross one box)

☐ did not go to school	☐ 15
☐ 13 or younger	☐ 16
☐ 14	☐ 17 or older

6. What qualification(s) do you have from school, college or the equivalent?

(please put a cross in the most appropriate box(es))

☐ clerical or commercial qualifications (eg secretarial, hairdressing etc)
☐ nursing or teaching
☐ "O" level (or equivalent)
☐ "A" level (or equivalent)
☐ college/university degree (or equivalent)
☐ none of these

7. About how many cigarettes do you smoke on average each day, now? (please cross one box)

☐ none	☐ 15-19
☐ less than 5	☐ 20-24
☐ 5-9	☐ 25 or more
☐ 10-14	

8. Are you an ex-smoker? ☐ No ☐ Yes

9. About how much wine, beer or spirits do you drink on average each week? (please cross one box for each type)

Wine (glasses per week)	Lager/Cider/Beer (half pints per week)	Spirits (tots per week)
☐ none	☐ none	☐ none
☐ less than 1	☐ less than 1	☐ less than 1
☐ 1-3	☐ 1-3	☐ 1-3
☐ 4-6	☐ 4-6	☐ 4-6
☐ 7-10	☐ 7-10	☐ 7-10
☐ 11-15	☐ 11-15	☐ 11-15
☐ 16-20	☐ 16-20	☐ 16-20
☐ 21+	☐ 21+	☐ 21+

If you drink wine is it

☐ mostly red	☐ mostly white
☐ about the same amount of red and white?	

10. How often do you do any exercise?

- ☐ rarely/never ☐ 2-3 times a week
☐ less than once a week ☐ 4-6 times a week
☐ once a week ☐ every day

11. How often do you do strenuous exercise?

(that is, enough to cause sweating or a fast heart beat.)

- ☐ rarely/never ☐ 2-3 times a week
☐ less than once a week ☐ 4-6 times a week
☐ once a week ☐ every day

QUESTIONS ABOUT YOU AND YOUR FAMILY

12. Have you ever had any children? ☐ No ☐ Yes

- if No, please go on to question 15

13. How many children have you had?

(please include stillbirths; it is not necessary to include miscarriages)

14. When was each child born, and for how many months did you breastfeed each child, if at all?

DATE OF BIRTH

(If you had twins or triplets please repeat the same date for each child)

day month year

1st child / /

2nd child / /

3rd child / /

4th child / /

5th child / /

6th child / /

7th child / /

8th child / /

9th child / /

10th child / /

BREASTFEEDING

(Months that you breastfed each child; put "0" if you did not breastfeed that child and "1" if you breastfed for a month or less)

months

months

months

months

months

months

months

months

months

months

15. Have you ever been for breast screening before?

- ☐ No
☐ Yes- If Yes, about how many years ago was your last screen? years ago

16. Have you ever had a breast lump removed or any operations on your breast(s)?

- ☐ No
☐ Yes- If Yes, how old were you? years

(If you have had more than one operation please write your age at the first operation)

17. Have you ever had breast cancer diagnosed?

- ☐ No
☐ Yes- If Yes, how old were you when the cancer was first diagnosed? years

18. Has your mother ever had breast cancer diagnosed?

- ☐ No ☐ Don't know
☐ Yes- If Yes, how old was she when the cancer was first diagnosed? years

19. How many sisters do you have? sisters

(put "0" if you do not have any sisters, please include any sisters who have died)

20. Have any of your sisters ever had breast cancer diagnosed?

- ☐ No/No sisters ☐ Don't know
☐ Yes- If Yes, how old were they when the cancer was first diagnosed?

1st sister years 2nd sister years

QUESTIONS ABOUT YOUR HEALTH

21. Have you ever had any other cancer?

- ☐ Yes ☐ No

Please describe

22. Have you EVER had:

(please cross "Yes" or "No" for each condition)

- High blood pressure - when pregnant ☐ Yes ☐ No
High blood pressure - when not pregnant ☐ Yes ☐ No
Heart disease(eg heart attack/angina) ☐ Yes ☐ No
Stroke ☐ Yes ☐ No
Diabetes ☐ Yes ☐ No
High blood cholesterol ☐ Yes ☐ No
Blood clot (thrombosis) ☐ Yes ☐ No

23. Are you NOW being treated for:

- High blood pressure (hypertension) ☐ Yes ☐ No
Heart disease ☐ Yes ☐ No
Diabetes ☐ Yes ☐ No
High blood cholesterol ☐ Yes ☐ No
Varicose veins ☐ Yes ☐ No
Clotting problems ☐ Yes ☐ No
Asthma ☐ Yes ☐ No
Rheumatoid arthritis ☐ Yes ☐ No
Osteoarthritis ☐ Yes ☐ No
Thyroid problems ☐ Yes ☐ No
Osteoporosis ☐ Yes ☐ No
Depression/Anxiety ☐ Yes ☐ No

24. Are you **NOW** being treated for any other *serious* illness?

☐ Yes ☐ No

Please describe this illness

Please describe the treatment

QUESTIONS ABOUT PAST OPERATIONS

25. Have you had a hysterectomy?

☐ No

☐ Yes- If **Yes**, how old were you? years

26. Have you had **BOTH** ovaries removed?

☐ No

☐ Not sure

☐ Yes- If **Yes**, how old were you? years

27. Have you been sterilised (had your tubes tied)?

☐ No

☐ Yes- If **Yes**, how old were you? years

QUESTIONS ABOUT YOUR USE OF THE PILL

28. Have you ever used the pill (oral contraceptive)?

☐ Yes

☐ No - if No, please go to question 32

29. About how old were you when you *first* went on the pill? years

30. About how old were you when you *last* came off the pill? years

31. For how many years *in total* did you take the pill?

years

(Add together the years and months when you actually took the pill -do not count the years and months when you were not taking it. Please write "0" if you used the pill for less than a year in total)

QUESTIONS ABOUT YOUR USE OF HORMONE REPLACEMENT THERAPY (HRT)

32. Have you ever used hormone replacement therapy (HRT)? ☐ No - if No - please go to question 39

☐ Yes

33. How old were you when you *first* started using HRT?

years

34. Had your periods stopped before you started using HRT? (Cross "Yes" if you had a hysterectomy before starting HRT)

☐ No

☐ Yes - if **Yes**, how old were you when your periods stopped? years

35. For about how many years *in total* have you used HRT? years

(Add together the years and months when you used HRT - do not count the years and months when you were not using HRT. Please write "0" if you used HRT for less than a year in total)

36. Are you *now* using HRT?

☐ Yes

☐ No - if No, how old were you when you last used HRT? years

37. What is the name of the most **RECENT** HRT you have used?

☐ Prempak C 0.625mg

☐ Premarin 0.625mg

☐ Prempak C 1.25mg

☐ Premarin 1.25mg

☐ Tridestra

☐ Evorel 25mcg/50mcg

☐ Trisequens

☐ Evorel 75mcg/100mcg

☐ Trisequens Forte

☐ Progynova 1mg

☐ Cycloprogynova 1mg

☐ Progynova 2mg

☐ Cycloprogynova 2mg

☐ Estraderm 25mcg

☐ Estrapak

☐ Estraderm 50mcg

☐ Estracombi

☐ Estraderm 100mcg

☐ Climaval 1mg

☐ Zumenon 1mg

☐ Climaval 2mg

☐ Zumenon 2mg

☐ Premique Cycle

☐ Ethinyloestradiol

☐ Premique

☐ Micronor

☐ Nuvelle

☐ Provera

☐ Kliofem

☐ Duphaston

☐ Livial

☐ Do not know ☐ Implants ☐ Oestrogen

Other (please write here)

38. For how many years *in total* did you use the most **recent** type of HRT? years

(Please write "0" if you used this recent HRT for less than a year in total)

QUESTIONS ABOUT YOUR PERIODS

39. About how old were you when your periods started? years

40. Have your periods NOW stopped?

Cross "Yes"-if you are not having periods now, either because of your menopause, after a hysterectomy or after stopping HRT.

Cross "No"- if you are still having regular periods now, even if they are because you are taking HRT.

Cross "Irregular"-if your periods have been irregular and you think it might be because of the menopause.

☐ Yes- If Yes, how old were you when they stopped? years

☐ No

☐ Irregular

FINAL SECTION

41. So that we can find out about your health in the future we may need to contact you again or look at your screening or medical records. We would be grateful if you gave us permission to contact you again or to use information from those records.

We guarantee that all information obtained will be treated with absolute confidentiality and used for medical research only. Of course, you do not have to give permission. Your response to this request will not affect your screening or the treatment you receive in any way.

If you give permission, please sign here and print your name, address and other details in the section below. Please print in BLOCK CAPITALS as clearly as possible.

Signature:

today's date:

Surname:

Given name(s):

House number and street:

District:

Town:

County:

Postcode:

Surname at birth:

Town of birth:

NHS (National Health Service) Number:

(The number on your medical card)

Breast Screening Number:

(This is in the top left hand corner of your **screening invitation letter** and starts with the letters LGL)

Please bring this questionnaire to your breast screening appointment

If you would like to post this questionnaire back to us, please send it to:

THE MILLION WOMEN STUDY CO-ORDINATING CENTRE
ICRF-CEU, GIBSON BUILDING,
RADCLIFFE INFIRMARY,
OXFORD OX2 6HE

FREEPHONE: 0800 262 872

Breast Screening Unit
Linton House
Thirlestaine Road
Cheltenham
GL53 7AS

LGL ☐

THANK YOU VERY MUCH FOR YOUR HELP

A.2 First repeat survey questionnaire (yellow)

THE MILLION WOMEN STUDY

Confidential National Study of Women's Health

The Million Women Study is an important national study of women's health. A few years ago you received the first questionnaire with your invitation to the National Health Service Breast Screening Programme. Your help is needed again. Can you find time to complete this second questionnaire? Some of the questions may seem familiar and others are new, but all will provide vital up to date information for the study. Your answers are valuable and important - the enclosed leaflet explains how the study will benefit women and improve medical knowledge world-wide. We very much hope you are still willing to be one of the Million Women in the study.

We guarantee that all information provided will be treated with absolute confidentiality and used for medical research only.

To help us read your answers please write as clearly as possible and complete the questionnaire as shown:

Please put a cross in the appropriate box(es) ☐ ☐

OR put numbers in the appropriate boxes

eg 28th October 1946 /

Any questions? Ring us on Freephone 0800 262 872

QUESTIONS ABOUT YOU AND YOUR HEALTH

Please answer every question as best you can as all the information that you give us is very useful. If you are not sure about exact dates or ages an approximate answer is better than none. Please use a black pen, if possible.

1. Have you had any of the following conditions diagnosed for the first time in the last 5 years?

Please cross "yes" if appropriate, and write when the condition was first diagnosed.

	Yes	If Yes, when was it first diagnosed?	
		Month	Year
Heart disease heart attack, angina etc	☐		
Stroke	☐		
Diabetes	☐		
Blood clot in leg	☐		
Blood clot in lung or elsewhere	☐		
High blood cholesterol	☐		
High blood pressure hypertension	☐		
Asthma	☐		
Gallstones/gallbladder disease	☐		
Osteoporosis	☐		
Abnormal cervical smear test Pap smear	☐		
Breast cancer	☐		
Other cancer Please describe the cancer below	☐		

2. Have you had any broken/fractured bones, in the last 5 years?

☐ No ☐ Yes - if Yes, when?

If Yes, which bone(s) were broken? (please cross)

☐ hip ☐ ankle ☐ wrist/arm
☐ spine ☐ other- please describe _____

If Yes, how did the fracture(s) occur?

☐ after a fall ☐ in a car accident ☐ in another accident
☐ fracture found on X-ray without you knowing about it ☐ in some other way

3. Have you had any other serious illness diagnosed for the first time, in the last 5 years?

☐ No ☐ Yes - if Yes, when?

If yes, please describe the illness

4. Have you had any major operations in the last 5 years?

☐ No ☐ Yes - if Yes, when?

If yes, please describe the operation(s)

5. Are you NOW being treated for: Yes

If yes, about how old were you when treatment began?

Heart disease heart attack, angina etc	☐		
Diabetes	☐		
High blood pressure hypertension	☐		
Asthma	☐		
Osteoarthritis	☐		
Osteoporosis	☐		
Any other serious illness or disability (please describe below)	☐		

6. How would you describe your health now?

☐ excellent ☐ good ☐ fair ☐ poor +

QUESTIONS ABOUT MEDICATIONS

7. Have you ever used HRT (hormone replacement therapy)? ☐ No - if No - please go to question 12 .

☐ Yes

8. How old were you when you first started using HRT? years old

9. Are you now using HRT?

☐ No - if No, when did you stop using HRT?

☐ Yes

Month Year

10. For about how many years in total have you used HRT? years of use
(Add together all the years and months when you were using HRT. Please write "0" if you used HRT for less than a year in total)

11. Which HRT did you use MOST RECENTLY?

- | | |
|---|--|
| ☐ Prempak C 0.625mg | ☐ Premarin 0.625mg |
| ☐ Prempak C 1.25mg | ☐ Premarin 1.25mg |
| ☐ Tridestra | ☐ Evorel 25mcg/50mcg |
| ☐ Trisequens | ☐ Evorel 75mcg/100mcg |
| ☐ Cycloprogynova 1mg | ☐ Progynova 1mg |
| ☐ Cycloprogynova 2mg | ☐ Progynova 2mg |
| ☐ Estrapak | ☐ Estraderm 25mcg |
| ☐ Estracombi | ☐ Estraderm 50mcg |
| ☐ Climaval 1mg | ☐ Estraderm 100mcg |
| ☐ Climaval 2mg | ☐ Zumenon 1mg |
| ☐ Premique Cycle | ☐ Zumenon 2mg |
| ☐ Premique | ☐ Ethinyloestradiol |
| ☐ Nuvelle | ☐ Oestrogel |
| ☐ Kliofem | ☐ Implants |
| ☐ Livial | ☐ Provera |
| | ☐ Micronor |
| | ☐ Duphaston |
| | ☐ Do not know |

Other (please write here)

12. Have you taken any medications (other than HRT) for most of the last 4 weeks? ☐ No ☐ Yes

If Yes, was it: ☐ thyroxine ☐ ibuprofen ☐ aspirin
☐ tamoxifen ☐ bendroflumazide ☐ amlodipine ☐ digoxin
☐ paracetamol ☐ propranolol ☐ atenolol ☐ warfarin
☐ prednisolone ☐ Losec/Zoton ☐ Prozac ☐ insulin
☐ co-proxamol ☐ amitriptyline ☐ sleeping pills ☐ lithium
☐ Distalgesic ☐ Tryptizol etc

Please give the name(s) of any other medication you have used for most of the last 4 weeks:

13. Do you regularly take any vitamins, minerals or supplements? ☐ No ☐ Yes - if Yes, do you take:

☐ multivitamins (with minerals) ☐ vitamin A
☐ multivitamins (without minerals) ☐ vitamin B (including B₆, B₁₂)
☐ fish oil (including cod liver oil) ☐ vitamin C
☐ evening primrose oil ☐ garlic ☐ vitamin D
☐ iron ☐ zinc ☐ calcium ☐ vitamin E

for office use only

QUESTIONS ABOUT YOUR DIET

We know it may be difficult for you to give exact answers to these questions about your diet. An approximate answer is very valuable for this study. So, please answer as best you can, thinking of a typical week.

14. Which types of meat do you eat about once a week or more often? (you can cross more than one box)

☐ beef	☐ bacon	☐ chicken/poultry
☐ lamb	☐ ham	☐ kidney
☐ pork	☐ sausages	☐ liver/pâté
☐ beefburger/hamburger	☐ never eat meat	

15. Which types of fish do you eat about once a week or more often? (you can cross more than one box)

☐ tuna	☐ trout	☐ mackerel	☐ "fish & chips"
☐ salmon	☐ sardines	☐ other seafood (prawns, scampi etc)	
☐ cod/haddock or other white fish	☐ never eat fish		

16. About how many times each week do you eat:

(please count all meals and snacks. put '0' if eaten less than once a week)

meat		number of times eaten each week (remember meat in sandwiches)
fish/seafood		number of times eaten each week
chips		number of times eaten each week
potatoes (except chips)		number of times eaten each week
pasta/spaghetti		number of times eaten each week
rice		number of times eaten each week
cheese		number of times eaten each week (remember cheese in pizzas, quiches, cheese sauce, etc)

17. About how many eggs do you eat each week?

eggs number of eggs eaten each week (remember eggs in omelettes, quiches, cakes etc. put '0' if less than one)

18. Which types of vegetables/salads (fresh, frozen or tinned) do you eat once a week or more often? (you can cross more than one box)

☐ green peas	☐ tomatoes	☐ green beans
☐ broccoli	☐ onions	☐ baked beans
☐ cabbage	☐ garlic	☐ soya meat/tofu
☐ carrots	☐ swede	☐ chick peas/lentils
☐ courgettes	☐ spinach	☐ cauliflower
☐ beetroot	☐ sweetcorn	☐ green/red peppers
☐ leeks	☐ avocado	☐ brussels sprouts
☐ parsnip	☐ aubergine	☐ mushrooms
☐ lettuce	☐ celery	☐ cucumber

19. About how much do you eat each week of:

(put "0" if less than one)

cooked vegetables (except potatoes)		number of heaped tablespoons each week
salad items/ raw vegetables		number of heaped tablespoons each week (please count lettuce, tomato etc in sandwiches)

20. Which types of fruit do you eat once a week or more often, when in season? (you can cross more than one box)

☐ apples	☐ bananas	☐ oranges, satsumas, etc
☐ grapefruit	☐ pears	☐ stone fruit (peaches, plums, nectarines etc)

21. About how much fruit or fruit juice do you eat or drink each week? (count 10 grapes, berries or raisins as one piece; put "0" if less than one a week)

number of pieces of fresh fruit eaten each week	number of pieces of dried fruit eaten each week
number of glasses of fruit juice each week	number of tablespoons of stewed or tinned fruit eaten each week

MORE ABOUT YOUR DIET

22. About how many of the following do you eat:

(put "0" if none or less than one)

slices/pieces of white bread		slices each week
slices/pieces of brown/wholemeal bread (also include granary, rye bread etc)		slices each week
crackers, crispbread etc (ryvita, water biscuits etc)		number each week
crisps, hula hoops etc		packets each week
sweet biscuits		number each week
dairy desserts (yoghurts etc)		number each week
cakes, puddings, pies, buns etc		number each week
chocolate (in any food or drink)		approx. number of pieces each week
boiled sweets, peppermints etc		number each week
nuts (including peanut butter)		tablespoons each week
gravy, cream/cheese sauces etc		tablespoons each week
jams, marmalade		tablespoons each week
breakfast type cereal		bowls each week

If you eat breakfast cereal is it usually: (please cross)

☐ bran cereal (allbran, branflakes etc)	☐ muesli
☐ biscuit cereal (weetabix, shreddies etc)	☐ other (cornflakes, rice crispies etc)
☐ oat cereal (porridge, ready brek etc)	

23. What type of spread do you use on bread, crispbreads etc, once a week or more often? (you can cross more than one box)

☐ butter	☐ margarine	☐ soft cheese
☐ low fat spread	☐ mayonnaise	☐ salad cream
☐ olive oil spread	☐ marmite etc	☐ rarely use spread

Do you spread it: ☐ thick? ☐ medium? ☐ thin? (please cross)

Do you add butter etc to: ☐ potatoes? ☐ other vegetables?

24. Which types of fats or oils do you use for cooking or salad dressing once a week or more often?

(you can cross more than one box)

☐ butter	☐ soft (tub) margarine	☐ white flora
☐ olive oil	☐ hard (block) margarine	☐ lard/dripping
☐ corn oil	☐ sunflower oil	☐ mayonnaise
☐ soya oil	☐ other vegetable oil	☐ salad cream

Please put a cross in the box if you RARELY OR NEVER:

☐ use fats or oils for cooking	☐ use salad dressing/cream
---	---

25. Please put a cross in the box if you NEVER eat:

☐ beef	☐ pork/ham	☐ lamb	☐ dairy products
☐ kidney	☐ liver/pâté	☐ sugar	☐ wheat products
☐ salami	☐ sausages	☐ eggs	☐ beefburgers

26. What type of milk or cream do you drink or use once a week or more often? (you can cross more than one box)

☐ full cream milk	☐ single cream
☐ semi-skimmed milk	☐ double cream
☐ skimmed/fat free milk	☐ dairy ice cream
☐ soya milk	☐ never have milk/cream

27. Do you:

never some-times usually always

add milk to your tea?	☐	☐	☐	☐
add milk to your coffee?	☐	☐	☐	☐
add salt to your food?	☐	☐	☐	☐
remove fat from meat? (cross "never" if vegetarian)	☐	☐	☐	☐
eat breakfast?	☐	☐	☐	☐
eat an afternoon snack?	☐	☐	☐	☐
eat organic food?	☐	☐	☐	☐

28. Have you made any major changes to your diet in the last 5 years? ☐ No ☐ Yes because of illness ☐ Yes for some other reason

29. About how much alcohol do you drink each week?

number of drinks of alcohol each week
(one drink = a glass of wine, half pint of lager, or tot of spirits; put "0" if you drink less than one drink each week)

If you have more than one drink of alcohol each week:

is it usually with meals? ☐ No ☐ Yes ☐ it varies

on how many days each week days each week
do you usually drink?

30. About how much do you drink EACH DAY of:

tea? cups daily		milk, hot chocolate etc? cups daily		fizzy/soft drink? glasses daily	
coffee? cups daily		water? glasses daily		fruit squash? glasses daily	

31. How many teaspoons of sugar do you add to tea,

coffee, cereal, fruit etc EACH DAY? teaspoons of sugar each day

32. What size clothes do you wear now?

(you can cross more than one box if the size varies)

Clothes	☐ 10 or less	☐ 12	☐ 14	☐ 16	☐ 18	☐ 20+
Bra	☐ 32	☐ 34	☐ 36	☐ 38	☐ 40	☐ 42+
Cup	☐ AAA	☐ B	☐ C	☐ D	☐ DD/E +	

33. What is your: (please put "0" if you do not know)

waist measurement?		inches	hip measurement?		inches
--------------------	----------------------	--------	------------------	----------------------	--------

34. About how much do you weigh now?

stone lbs (Put "0" if you do not know)

WHEN YOU WERE YOUNG

35. About how much did you weigh when you were born?

lbs ozs (Put "0" if you do not know)

36. Were you breastfed when you were a baby?

☐ No ☐ Yes ☐ do not know

37. Did your parents smoke at around the time that you were born?

Mother	☐ No	☐ Yes	☐ do not know
Father	☐ No	☐ Yes	☐ do not know

38. Did your parents smoke at around the time that you were 10 years old?

Mother	☐ No	☐ Yes	☐ do not know
Father	☐ No	☐ Yes	☐ do not know

39. When you were about 10 years old, compared to average, would you describe yourself as (please cross):

☐ thinner? ☐ plumper? ☐ about average?

40. What size clothes did you wear when you were about 20 years old? (you can cross more than one box)

☐ 8 or less ☐ 10 ☐ 12 ☐ 14 ☐ 16 ☐ 18 +

QUESTIONS ABOUT YOUR FAMILY AND LIFESTYLE

41. Is your mother still alive?

- ☐ Yes-please give her age now years old
- ☐ No-please give her age when she died years old
- ☐ Do not know

42. If your mother has died, what did she die from?

- ☐ heart disease ☐ breast cancer
heart attack etc
- ☐ stroke ☐ cancer of the womb
- ☐ chest infection ☐ cancer of ovary
pneumonia
- ☐ "old age" ☐ other/unknown _____

43. Has your mother or father ever suffered from:

- | | | | |
|--------------------------|--------------------------|--------------------------|--------------------------|
| mother | father | mother | father |
| ☐ | ☐ | ☐ | ☐ |
| ☐ | ☐ | ☐ | ☐ |
| ☐ | ☐ | ☐ | ☐ |
| ☐ | ☐ | ☐ | ☐ |
| ☐ | ☐ | ☐ | ☐ |
| ☐ | ☐ | ☐ | ☐ |
| ☐ | ☐ | ☐ | ☐ |
| ☐ | ☐ | ☐ | ☐ |
| ☐ | ☐ | ☐ | ☐ |

44. How tall is/was your mother? feet inches (Put "0" if you do not know)

How tall is/was your father? feet inches

45. Have you ever been a smoker?

- ☐ No - **if No** - please go to question 49.
- ☐ Yes

46. How old were you when you started smoking regularly? years old

47. Are you a smoker now?

- ☐ No - **if No** - how old were you when you stopped smoking? years old
- ☐ Yes - **if Yes** - please write the tar & nicotine content of your usual brand of cigarettes:
(this is written on each packet of cigarettes) tar mg nicotine mg

48. About how many cigarettes do you/did you smoke on average each day? (If you are an ex-smoker, how many did you smoke on average when you smoked?)

cigarettes per day

49. Have you had your menopause?

- ☐ No ☐ Not sure (because periods irregular, taking HRT etc)
- ☐ Yes- How old were you when you had your menopause? years old

50. Are you now in paid work?

- ☐ No ☐ Yes, full time ☐ Yes, part time

51. Are you currently married or living with a partner?

- ☐ No ☐ Yes- **if Yes**- does your husband/partner smoke? ☐ No ☐ Yes

52. About how often do you use a mobile phone?

- ☐ Never ☐ less than once a day ☐ every day

For how long have you used one? years
(put "0" if never or less than 1 year)

53. Do you belong to or participate in any of the following?

- ☐ religious group ☐ art/craft group ☐ bingo
- ☐ voluntary work ☐ music/singing group
- ☐ adult education ☐ sports club (swimming, golf etc)
- ☐ dancing group ☐ yoga, etc ☐ other group activity

54. How often do you feel:

	rarely/ never	sometimes	usually	most of the time
happy	☐	☐	☐	☐
relaxed	☐	☐	☐	☐
in control	☐	☐	☐	☐
stressed	☐	☐	☐	☐
tired	☐	☐	☐	☐

55. Do you have a nap during the day?

- ☐ rarely/never ☐ sometimes ☐ usually

56. About how many hours sleep do you get in every 24 hours? hours sleep (please include naps)

57. To which ethnic group do you consider you belong?

- ☐ White ☐ Black - Caribbean, African etc.
- ☐ Asian ☐ Other -please specify _____

58. What is your date of birth?

Day	Month	Year

59. On what date did you fill in this form?

----------------------	----------------------	----------------------	----------------------

60. In case we need to check on any details, it would be helpful if you would write your telephone number below.

STD code										Telephone number									

THANK YOU VERY MUCH FOR YOUR HELP

Please put your completed questionnaire in the pre-paid envelope and post it back to us

If your name /address has changed or is incorrect could you please cross this box ☐ & give the correct details below.

Surname:

Given name(s):

House number and street:

District:

Town/Country: Postcode:

For office use only			
☐	☐	☐	☐
☐	☐	☐	☐
☐	☐	☐	☐

MWS-PF/102
9909

PROFESSOR VALERIE BERL, THE MILLION WOMEN STUDY, ICRF-CEU, GIBSON BUILDING,
 RADCLIFFE INFIRMARY, OXFORD OX2 6HE
 FREEPHONE: 0800 262 872

Appendix B

Reliability of self-reported childbearing data in the Million Women Study

Abstract

Objectives: *To assess how reliably women recall their childbearing histories (parity and breastfeeding)*

Design: *Comparison of self-reported childbearing histories*

Setting: *Participants from the Million Women Study who completed a recruitment questionnaire on more than one occasion*

Participants: *19,059 women who had inadvertently been recruited more than once, and who completed and returned more than one recruitment questionnaire*

Main outcome measure: *Degree to which childbearing data on the two questionnaires agrees*

Results: *19,059 (1.5%) women sent back more than one completed recruitment questionnaire. The mean time between questionnaires was two years. Women reliably reported their childbearing histories. For example, 99% of women who reported being parous on their initial recruitment questionnaire reported being parous on the second recruitment questionnaire they completed (Spearman's correlation 0.98 $p < 0.0000$.) The correlation between reported parity levels was also very good (Spearman's correlation 0.94 $P < 0.0000$.) Among parous women who reported having ever breastfed on their initial questionnaire 96% reported having breastfed on the subsequent questionnaire. Similarly among women who reported never having breastfed 94% on the initial questionnaire reported never having breastfed subsequently (Spearman's 0.87 $p < 0.0000$.) Women who breastfed reliably reported the duration they had breastfed. The mean duration of breastfeeding in this group of women was 7.1 months (SD 8.6) on the first questionnaire and 7.51 months (SD 8.7) on the second questionnaire. 65% of women reported the same duration on both occasions and over 90% of women reported breastfeeding durations within 3 months of each other. The Spearman's correlation for duration of breastfeeding was 0.90 $p < 0.0000$. A Bland-Altman plot fitted to duration of breastfeeding data suggested there was no systematic under or over-reporting of breastfeeding duration (mean difference of -0.12 months; 95% -7.4 to 7.1 months.)*

Conclusions: *In a sub-set of women who completed more than one recruitment questionnaire about two years apart childbearing histories were reliably reported. There was no evidence to suggest duration of breastfeeding data were substantially affected by systematic biases. These findings add weight to the existing evidence that childbearing data used in the Million Women Study is reliably reported by participants.*

Appendix C

Cardiovascular Risk in black South Africans study (CRiBSA)

Included is a summary version of the CRiBSA study protocol provided with generous permission of Professor N.Levitt

C.1 Summary study protocol: Cardiovascular Risk in black South

Africans study (CRiBSA)

PROTOCOL: The prevalence of diabetes mellitus and cardiovascular disease risk factors in urban black South Africans

Research Institutions: Department of Medicine, University of Cape Town, South Africa
Chronic Diseases of Lifestyle Unit, Medical Research Council, SA
Sports Science Institute

Principal Investigator: Professor Naomi S Levitt

Collaborators: Professor Krisela Steyn, Professor Estelle Lambert, Dr Nelia Steyn, Dr Nasheeta Peer, Dr Carl Lombard

BRIEF OUTLINE OF THE OVERALL PROTOCOL

Overall Aim

To describe and compare the prevalence of diabetes and cardiovascular disease risk (CV) factors in urban black South Africans.

Specific Objectives

1. To identify the major CVD risk factors in urban black South Africans.
2. To quantify the degree of risk associated with the various risk factors.
3. To provide the Department of Health with essential health data required to make health policy for a population in transition

Rationale

The main rationale for this study is the increase in development associated with urbanisation and industrialisation in South Africa with the resultant adoption of western habits such as high fat, high sodium, low fibre diet, increasing smoking and alcohol consumption and decreasing traditional diets and physical activity. However, there are no recent prevalence studies of diabetes and other CVD risk factors indicating current levels. The last comprehensive surveys of diabetes and CVD risk factors were conducted in the early 1990s in South Africa

There has been inadequate attention paid to the rising burden of chronic diseases in developing countries. This increased chronic disease burden together with the already high impact of infectious diseases including HIV/AIDS is straining health services in developing countries. In South Africa, the years of life lost for chronic diseases (25%) is close approaching that for infectious diseases (28%) (Yach et al, 2004: 2617). Therefore, despite its quadruple burden of disease that include poverty related diseases, emerging chronic diseases, injuries and HIV/AIDS, South Africa cannot afford to lose sight of chronic diseases that are a significant contributor to mortality and morbidity in the country.

The latest mortality statistics for South Africa indicate that cardiovascular diseases and diabetes mellitus are amongst the causes of death (Reddy, 2004). Although the South African National Demographic and Health survey provide excellent data on the prevalence of hypertension and smoking, it is unable to answer the question of the extent of diabetes or of the blood lipid profile abnormalities.

Proposed Methodology

Study Design

A cross sectional descriptive, analytical design will be used in the selected urban areas.

Study Population

The study population will consist of randomly selected ambulant non-pregnant residents between the ages of 25 and 65 in the traditionally urban black African townships of the Cape Town metropolitan area.

Sample Size and Frame

The sample size will be calculated based on an expected diabetes prevalence of 8% (Alberts et al, 2005 and Levitt et al, 1993), providing the justification for n=1 000 each, in both the urban and rural areas. A 95% confidence interval powered at 90% will be used to calculate the population parameters.

The sampling frame will be obtained from the South African Demographic and Health Survey which used enumeration areas (EA) created by Statistics South Africa. The EAs range from 100 to 250 households with the number of households in the EA serving as a measure of the size of the EA.

Participants will be selected on a quota for age and gender categorisation; the rationale being that the risk factors and disease prevalence differ among age groups and genders. There will be 100 to 200 participants in each of the 10 strata, stratified for age (25 – 34, 35 – 44, 45 – 54 and 55 – 64) and gender (male and female).

Data Collection and Instruments

The following will be examined in order to assess risk factors:

1. Dietary intake and habits: This will be collected by means of three 24-hour recalls conducted by trained nutritionists/dietitians. From this data the researchers will be able to analyse: energy and macronutrient intakes, micronutrient intakes, sodium, fatty acids and cholesterol intakes, antioxidants, fruit and vegetable intake and fibre intake (sodium will be measured by 24-hour urine collections).
2. Lipid profile: A fasting blood sample will be taken to ascertain total cholesterol, HDL cholesterol, LDL cholesterol, triglycerides, omega-3 and omega-6 fatty acids. Various micronutrients will also be analysed to validate dietary data.
3. Weight status: Weight, height, waist and hip circumferences will be measured using standardised methods. Percent body fat will be determined from skin folds and body impedance. Overweight and central obesity will be assessed based on Body Mass Index (BMI), waist and Waist Hip Ratio (WHR).
4. Blood pressure will be measured three times at five minute intervals on the subject's left arm while seated and at rest, using an automated technique. The appropriate cuff size will be used with a large cuff for subjects with an arm circumference ≥ 33 cm.
5. Insulin resistance (IR), impaired glucose tolerance (IGT) and fasting glucose will be measured. A standard Oral Glucose Tolerance Test (OGTT) will be administered: Each participant will undergo a standard 75g OGTT with blood drawn at 0 and 120 min. All samples will be stored and assayed at the University of Cape Town laboratories.
6. Alcohol intake will be assessed using a validated questionnaire (CAGE) and supported by blood data.
7. Tobacco-use will be assessed by means of a pretested questionnaire which has been validated in local populations. Subjects will also be tested with cotinine to support findings from the questionnaire.
8. Physical fitness will be tested using Eurofit testing battery for which there are South African norms. Physical activity habits will be assessed using a standardised questionnaire.
9. Psychosocial Parameters related to Diabetes will be estimated using questionnaires that have previously been used in this population.

Proposed Plan of Analysis

Descriptive statistics will be used to describe the population characteristics and prevalences. Bivariate and multivariate analyses will be used to explore associations.

The prevalence of diabetes, IR and IGT will be estimated and these will be modelled on various risk factors. Estimations will be done within sex and age strata. The association of age with some of the biochemical variables will be of interest and polynomial regression will be used to assess whether an association is linear or non-linear. Further analyses of the data will be planned according to the specific needs of the researchers in consultation with the statisticians responsible for the analyses.

Key References

1. Alberts M, Urdal P, Steyn K, Stensvold I, Tverdal A, Nel JH, Steyn NP. Prevalence of common non-communicable diseases and associated risk factors in a rural black population of South Africa. *European Journal of Cardiovascular Prevention and Rehabilitation* 2005, 12:347-54
2. Levitt NS, Katzellenbogen JM, Bradshaw D, Hoffman MN, Bonnici F. The prevalence and identification of risk factors for NIDDM in urban Africans in Cape Town, South Africa. *Diabetes Care* 1993, 16(4): 601-7
3. Mollentze WF, Moore AJ, Steyn AF, Joubert G, Steyn K, Oosthuizen GM, Weich DJV. Coronary heart disease risk factors in a rural and urban Orange Free State population. *SAMJ* 1995, 85: 90-6
4. Rosmond R. Role of stress in the pathogenesis of the metabolic syndrome. *Psychoneuroendocrinology* 2005, 30: 1-10
5. Reddy P. Chronic diseases, South African Health Review 2003/04, *Health Systems Trust* 2004, 175-87
6. Yach D, Hawkes C, Gould CL, Hofman KJ. The global burden of chronic diseases overcoming impediments to prevention and control. *JAMA* 2004; 291: 2616-22

Appendix D

Questionnaires and training materials for childbearing patterns in and health profiles of a population of middle-aged black South African women

D.1 English-language version of questionnaire

Cardio-metabolic disease & reproductive histories questionnaire – CRiBSA 2008

Interviewer	For Office Use ONLY
<u>Section 1: General Information</u>	
Study Number	
Date of interview.....	
Name of interviewer.....	
1A. Current Age.....	
1B. Date of Birth.....	
<u>Section 2: Menstrual History</u>	
1 = YES 2 = NO 9 = I don't know	
2 A. How old were you when you had your first period?	
[Please fill in YEAR if unable to recall age]	
2 B. Are you still having REGULAR menstrual periods?	
PLEASE NOTE If irregular or episodes of post-menopausal bleeding	

If NO,	
2 C. How old were when you had your last menstrual period?	
[Please fill in YEAR if unable to recall age]	
<u>Section 3: Pregnancy & Birth information</u>	
1 = YES 2 = NO 9 = I don't know	
3 A. Were you ever pregnant?	
If YES,	
3 B. How many times have you been pregnant ?	
[Any pregnancies including any losses that may have occurred]	
If NO please go to questions on CONTRACEPTIVE use	
3 C. How many children did you deliver/give birth to?	
[Please ONLY include live or still births NOT other pregnancy losses]	
3 D. How many live births?	
3 E. How many still births?	
3 F. How many children of your own did you breastfeed for 1 month of more?	
3 G. did you ever breastfeed children that were not your own (including breast milk donations)?	
If YES please COMPLETE Section 6	

Cardio-metabolic disease & reproductive histories questionnaire – CRiBSA 2008

3 H. Interviewer – fill in YEAR of birth if can't recall full date of birth. PLEASE ONLY INCLUDE LIVE OR STILL BIRTHS. PLEASE COMPLETE BREASTFEEDING DURATION INFORMATION Please note any pregnancy over 6 months which was lost to be recorded as <u>still birth</u> Please record <u>still birth as 999 in Reason for Stopping</u>													For Office use ONLY			
Child	Date of birth of child								Duration Breastfeeding		Reason for stopping			YoB Duration [Months]		
	d	d	m	m	y	y	y	y	Months							
								Years								
Child	Date of birth of child								Duration Breastfeeding		Reason for stopping			YoB Duration [Months]		
	d	d	m	m	y	y	y	y	Months							
								Years								
Child	Date of birth of child								Duration Breastfeeding		Reason for stopping			YoB Duration [Months]		
	d	d	m	m	y	y	y	y	Months							
								Years								
Child	Date of birth of child								Duration Breastfeeding		Reason for stopping			YoB Duration [Months]		
	d	d	m	m	y	y	y	y	Months							
								Years								
Child	Date of birth of child								Duration Breastfeeding		Reason for stopping			YoB Duration [Months]		
	d	d	m	m	y	y	y	y	Months							
								Years								
Child	Date of birth of child								Duration Breastfeeding		Reason for stopping			YoB Duration [Months]		
	d	d	m	m	y	y	y	y	Months							
								Years								
Child	Date of birth of child								Duration Breastfeeding		Reason for stopping			YoB Duration [Months]		
	d	d	m	m	y	y	y	y	Months							
								Years								
Child	Date of birth of child								Duration Breastfeeding		Reason for stopping			YoB Duration [Months]		
	d	d	m	m	y	y	y	y	Months							
								Years								
Child	Date of birth of child								Duration Breastfeeding		Reason for stopping			YoB Duration [Months]		
	d	d	m	m	y	y	y	y	Months							
								Years								
Child	Date of birth of child								Duration Breastfeeding		Reason for stopping			YoB Duration [Months]		
	d	d	m	m	y	y	y	y	Months							
								Years								

Cardio-metabolic disease & reproductive histories questionnaire – CRiBSA 2008

3F. How old were you when you delivered your FIRST child? [Please fill in YEAR if unable to recall age] (If required fill in from detailed birth history table) 3.G How old were you when you delivered your LAST child ? [Please fill in YEAR if unable to recall age] (If required fill in from detailed birth history table) How old were you when you breast fed your LAST child? [Please fill in YEAR if unable to recall age] (If required fill in from detailed birth history table) 3 H. During any of your pregnancies did you suffer from High blood pressure? Diabetes?	
--	---

Section 4: Use of hormonal contraceptives or HRT

Interviewer	For Office use ONLY
1 = YES 2 = NO 9 = I don't know	
4 A. Did you ever use oral contraceptives (the "pill")?	
If YES,	
4 B. Are you still using oral contraceptives?	
4 C. Did you ever use injectable contraception? ("the injection"/"depo"/"nur-isterate"/"petogen")?	
If YES,	
4 D. Are you currently using injectable contraception?	
4 E. Have you ever used hormone replacement therapy ("hormone therapy"/"HRT")?	
If YES,,	
4 F. Are you currently using hormone replacement therapy?	

Cardio-metabolic disease & reproductive histories questionnaire – CRiBSA 2008

Section 5: Additional Breastfeeding Information

Interviewer				For Office use ONLY			
For women who breastfed other children, not their own							
Child	Year OR Age when breastfeeding			Duration breastfeeding			
				Months			Age
				Years			BF Duration [Months]
				Months			Age
				Years			BF Duration [Months]
				Months			Age
				Years			BF Duration [Months]

Total for Sections 4 & 5

For Office use ONLY			
Summary			
Total Children Born			
Total Children BF			
Total life time duration bf			

Completed

For Office use ONLY			
Section 1	Study Number		
Section 2	Dates OR Years		
Section 3	Age or YEAR		
	Pregnancies		
	Births		
	Birth Order Correct		
	G.P.		
	Number Breastfed		
	Total BF Duration (M)		
Section 4	USE		
Section 5	Age or YEAR		

D.2 isiXhosa-language version of questionnaire

Cardio-metabolic disease & reproductive histories questionnaire – CRiBSA 2008

Interviewer	For Office Use ONLY	
<u>Section 1: General Information</u>		
Study Number 		
Date of interview.....		
Name of interviewer.....		
1 A. Iminyaka yakho ngoku.....		
1 B. Umhla wokuzalwa.....		
<u>Section 2: Menstrual History</u>		
1 = YES /EWE 2 = NO/HAYI 9 = I don't know/Andiyazi		
2 A. Waqala uneminyaka emingaphi ukuya exesheni? [Please fill in YEAR if unable to recall age] 		
2 B. Are you still having REGULAR menstrual periods? PLEASE NOTE If irregular or episodes of post-menopausal bleeding		
Ukuba ngu-hayi,		
2 C. How old were when you had your last menstrual period? [Please fill in YEAR if unable to recall age] 		

<u>Section 3: Pregnancy & Birth information</u>		
1 = YES /EWE 2 = NO/HAYI 9 = I don't know/Andiyazi		
3 A. Wakhe wakhulelwa kwixesha elingaphambili ? 		
Ukuba ngu-ewe,		
3 B. Ukhulelwe amatyeli amangaphi kuquka umntwana ovele ephila kwakunye nomntwana ovele engasaphili ? 		
PLEASE ONLY INCLUDE LIVE OR STILL BIRTHS. NOT OTHER LOSES Please note any pregnancy over 6 months which was lost to be recorded as <u>still birth</u>		
If NO please go to questions on CONTRACEPTIVE use		
3 C. Bangaphi abantwana obazeleyo ? [Nceda ungabali namasu aphumileyo] 		
3 D. Bangaphi abazelwe bephila? 		
3 E. Bangaphi abazelwe bengasaphili? [Still Birth] 		
3 F. Kubantwana obazeleyo, bangaphi kubo obancancise isithuba esingangenyanga nangaphezulu ? 		
3 G. Wakhe wancancisa abantwana abangengabakho ? (including breast milk donations)? 		
If YES please COMPLETE Section 6		

Cardio-metabolic disease & reproductive histories questionnaire – CRiBSA 2008

3 H. Interviewer – fill in YEAR of birth if can't recall full date of birth. PLEASE ONLY INCLUDE LIVE OR STILL BIRTHS. PLEASE COMPLETE BREASTFEEDING DURATION INFORMATION Please note any pregnancy over 6 months which was lost to be recorded as <u>still birth</u> Please record <u>still birth as 999 in Reason for Stopping</u>													For Office use ONLY													
Child	Date of birth of child								Duration Breastfeeding			Reason for stopping		YoB	<table border="1"> <tr><td></td><td></td><td></td><td></td></tr> <tr><td></td><td></td><td></td><td></td></tr> </table>											
	d	d	m	m	y	y	y	y	Months																	
Years																										
Child	Date of birth of child								Duration Breastfeeding			Reason for stopping		YoB	<table border="1"> <tr><td></td><td></td><td></td><td></td></tr> <tr><td></td><td></td><td></td><td></td></tr> </table>											
	d	d	m	m	y	y	y	y	Months																	
Years																										
Child	Date of birth of child								Duration Breastfeeding			Reason for stopping		YoB	<table border="1"> <tr><td></td><td></td><td></td><td></td></tr> <tr><td></td><td></td><td></td><td></td></tr> </table>											
	d	d	m	m	y	y	y	y	Months																	
Years																										
Child	Date of birth of child								Duration Breastfeeding			Reason for stopping		YoB	<table border="1"> <tr><td></td><td></td><td></td><td></td></tr> <tr><td></td><td></td><td></td><td></td></tr> </table>											
	d	d	m	m	y	y	y	y	Months																	
Years																										
Child	Date of birth of child								Duration Breastfeeding			Reason for stopping		YoB	<table border="1"> <tr><td></td><td></td><td></td><td></td></tr> <tr><td></td><td></td><td></td><td></td></tr> </table>											
	d	d	m	m	y	y	y	y	Months																	
Years																										
Child	Date of birth of child								Duration Breastfeeding			Reason for stopping		YoB	<table border="1"> <tr><td></td><td></td><td></td><td></td></tr> <tr><td></td><td></td><td></td><td></td></tr> </table>											
	d	d	m	m	y	y	y	y	Months																	
Years																										
Child	Date of birth of child								Duration Breastfeeding			Reason for stopping		YoB	<table border="1"> <tr><td></td><td></td><td></td><td></td></tr> <tr><td></td><td></td><td></td><td></td></tr> </table>											
	d	d	m	m	y	y	y	y	Months																	
Years																										
Child	Date of birth of child								Duration Breastfeeding			Reason for stopping		YoB	<table border="1"> <tr><td></td><td></td><td></td><td></td></tr> <tr><td></td><td></td><td></td><td></td></tr> </table>											
	d	d	m	m	y	y	y	y	Months																	
Years																										
Child	Date of birth of child								Duration Breastfeeding			Reason for stopping		YoB	<table border="1"> <tr><td></td><td></td><td></td><td></td></tr> <tr><td></td><td></td><td></td><td></td></tr> </table>											
	d	d	m	m	y	y	y	y	Months																	
Years																										

Cardio-metabolic disease & reproductive histories questionnaire – CRiBSA 2008

3 H. Umfumene uneminyaka emingaphi umntwana wakho wokuqala (ozelwe ephila okanye ozelwe engaphili)?	<table border="1"><tr><td></td><td></td><td></td><td></td></tr><tr><td></td><td></td><td></td><td></td></tr></table>									<table border="1"><tr><td></td><td></td><td></td><td></td></tr><tr><td></td><td></td><td></td><td></td></tr></table>								
[Please fill in YEAR if unable to recall age]																		
(If required fill in from detailed birth history table)																		
3 I. Ubuneminyaka emingaphi ukugqibela kwakho ukuzala ?	<table border="1"><tr><td></td><td></td><td></td><td></td></tr><tr><td></td><td></td><td></td><td></td></tr></table>									<table border="1"><tr><td></td><td></td><td></td><td></td></tr><tr><td></td><td></td><td></td><td></td></tr></table>								
[Please fill in YEAR if unable to recall age]																		
(If required fill in from detailed birth history table)																		
3 J. How old were you when you breast fed your LAST child?	<table border="1"><tr><td></td><td></td><td></td><td></td></tr><tr><td></td><td></td><td></td><td></td></tr></table>									<table border="1"><tr><td></td><td></td><td></td><td></td></tr><tr><td></td><td></td><td></td><td></td></tr></table>								
[Please fill in YEAR if unable to recall age]																		
(If required fill in from detailed birth history table- round up)																		
3 K. Xa uthe wakhulelwa uye ubenazo na iingxaki zokonyukelwa																		
Luxinzelelo-gazi oluphezulu?	<table border="1"><tr><td></td><td></td></tr><tr><td></td><td></td></tr></table>					<table border="1"><tr><td></td><td></td></tr><tr><td></td><td></td></tr></table>												
Isifo seswekile ngethuba ukhulelwe?	<table border="1"><tr><td></td><td></td></tr><tr><td></td><td></td></tr></table>					<table border="1"><tr><td></td><td></td></tr><tr><td></td><td></td></tr></table>												

Section 4: Use of hormonal contraceptives or HRT

Interviewer	For Office use ONLY
1 = YES /EWE 2 = NO/HAYI 9 = I don't know/Andiyazi	
4 A. Wakhe wasebenzisa iipilisi zokucwangcisa (“iipilisi”)? Ukuba ngu-ewe,	
4 B. Ingaba usebenzisa iipilisi zokucwangcisa nangoku?	
4 C. Wakhe wasebenzisa inaliti yokucwangcisa (“inalitit”/”idepo”/”nur-isterate”/”petogen”)? Ukuba ngu-ewe,	
4 D. Ingaba usebenzisa inaliti yokucwangcisa nangoku?	
4 E. Wakhe wasebenzisa ichiza lokogneza incindi yedlala (“hormone therapy”)? Ukuba ngu-ewe,	
4 F. Ingaba usebenzisa ichiza lokongeza incindi yedlala nangoku?	

Cardio-metabolic disease & reproductive histories questionnaire – CRiBSA 2008

Section 5: Additional Breastfeeding Information

Interviewer				For Office use ONLY			
For women who breastfed other children, not their own							
Child	Year OR Age when breastfeeding			Duration breastfeeding			
				Months			Age
				Years			BF Duration [Months]
				Months			Age
				Years			BF Duration [Months]
				Months			Age
				Years			BF Duration [Months]

Total for Sections 4 & 5

For Office use ONLY			
Summary			
Total Children Born			
Total Children BF			
Total life time duration bf			

Completed	For Office use ONLY			
	Section 1	Study Number		
	Section 2	Dates OR Years		
	Section 3	Age or YEAR		
		Pregnancies		
		Births		
		Birth Order Correct		
		G.P.		
		Number Breastfed		
		Total BF Duration (M)		
Section 4	USE			
Section 5	Age or YEAR			

D.3 English-language written training materials

Body mass indexes & reproductive histories of middle-age black South African women

Project:

A parallel study attached to the CRiBSA survey exploring women's reproductive histories in relation to their body mass index (BMI) in middle-age.

Research Aims:

- Accurately describe reproductive histories (including parity and breastfeeding) in middle-age black South African women
- Explore how parity (the number of full term pregnancies) and breastfeeding affect women's body mass index (BMI) in middle-age

Notes on the questionnaire:

The reproductive history questionnaire explores many different aspects of women's reproductive histories; from when they start their periods (menstruation) until they complete their families and experience "change of life" (the menopause).

The questionnaire is available in both IsiXhosa and English. It should be administered to every participant over the age of 40 years. It contains five sections spread over four pages.

Each page is divided into two parts. The left-hand side is to be completed by the interviewer, the right-hand side is used for subsequent data coding [see figure 1]. Most of the questions are "yes/no" questions. Interviewers should please record 1 for "YES" and 2 for "NO" in the appropriate blocks. If a woman does not know the answer or can't remember please record a 9 in the space provided. Additional space is provided for any other notes the interviewer would like to make.

Figure 1. Example of section 5 from the English version of the questionnaire

Section 5: Additional Breastfeeding Information					
Interviewer				For Office use ONLY	
For women who breastfed other children, not their own					
Child	Year OR Age when breastfeeding	Duration breastfeeding			
		Months		Age	
		Years		BF Duration [Months]	
		Months		Age	
		Years		BF Duration [Months]	
		Months		Age	
		Years		BF Duration [Months]	

Additional notes by section:

Section 1 – General participant information

This includes study number, current age and date of birth. We ask about both age and date of birth so we can double check this information. Don't worry if a woman can only remember the year she was born, record as much information as possible.

Section 2 – Menstrual history

Most women start having menstrual cycles (periods) some time between the ages of 9 and 16. Women may have regular or irregular periods during their lifetime. Periods stop during pregnancy and for a variable amount of time after giving birth. Women who use injectable contraceptives (depo/petogen/"the injection") may also stop having periods; these usually return after women stop using this type of contraception though they may take some time before returning to "normal".

Women usually experience "change of life" or menopause between the ages of 45 and 55. Usually a woman's periods will become irregular with longer intervals between each cycle until they stop completely. Women who describe vaginal bleeding after they have completed their menopause have "post-menopausal bleeding"; this is NOT normal and may be a sign of illness. If a woman reports post-menopausal bleeding or the interviewer is unsure they should ask for assistance from Serena or Debbie. If required the woman should be referred for a gynaecological evaluation.

Section 3 – Pregnancy & Birth Information

This section is one of the most important as its main focus is on the childbearing histories and experiences of our female participants. We need as much information as possible including any extra notes interviewers wish to make.

The medical term for when a woman becomes pregnant is gravidity. When doctor's report on the number of times a woman has become pregnant they will often use the term "grav" or simply "G" followed by the number. Parity is the term used to describe a birth. Thus a woman is described as a G3P2 it means she has been pregnant three times, but only given birth twice [this may be because she is pregnant now, or because she "lost" a pregnancy]. If a woman is described as G3P4 it means she has been pregnant three times but has given birth to four children, this would happen if she had twins for example.

If a woman says she never fell pregnant please record 2 for NO in answer to question 3A and go directly to section 4 which asks about her use of contraceptives.

With each pregnancy there is a good chance that the woman will deliver a "full-term" baby- that is a baby that is fully developed [usually between 38-42 weeks of gestation.] Sometimes, however, pregnancies end early. Some of the reasons for this include spontaneous miscarriage, pregnancy termination, or a pregnancy loss.

In this study a miscarriage is defined as a pregnancy ends before the first six months [24 weeks] – please DO NOT record these in this questionnaire. If a woman loses a pregnancy after six months we describe a "pregnancy loss" and record this as a birth. When a child is born it is recorded either as a live birth or a still birth. A still birth is defined in this study as a child who is born dead, if the child is born alive [breathes after birth] then this is recorded as a live birth. This applies to any pregnancy that lasted longer than 6 months.

Section 3H asks for specific information about each pregnancy longer than six months, the number of children born, the length of time each child was breastfed in months, and reasons for stopping breastfeeding.

There is space for up to 10 births. If a woman has given birth to more than ten children please record the first ten births and make a note of the total number of births [remember only those pregnancies that were longer than six months should be included, so don't record pregnancies that ended in miscarriage]. For each birth there is room for the date of birth [it is fine if a woman can only recall the year of birth, if she is unable please use the code 9999]. For multiple pregnancies, twins or more please record all the children even though they will have the same date of birth.

For each birth women should be asked if they breastfed. If they did not please record 00 in the Duration of Breastfeeding blocks. Please also record the reason they did not breastfeed. If the reason is because they had a still birth please record the code 999. If a woman reports breastfeeding please report the length of time she breastfed in months and years in the blocks provided. For example if a woman reports she breastfed a child for a year and a half this would be recorded as 01 in the year blocks and 06 in the month blocks. There is also space to make a note of why breastfeeding was stopped. This should be asked as an open ended question "can you remember why you stopped breastfeeding this child?"

Figure 2. Example from section 3H

3 H. Interviewer – fill in YEAR of birth if can't recall full date of birth. PLEASE ONLY INCLUDE LIVE OR STILL BIRTHS. PLEASE COMPLETE BREASTFEEDING DURATION INFORMATION Please note any pregnancy over 6 months which was lost to be recorded as <u>still birth</u> Please record <u>still birth as 999 in Reason for Stopping</u>													For Office use ONLY																
Child	Date of birth of child								Duration Breastfeeding			Reason for stopping	YoB <table border="1" style="display: inline-table; vertical-align: middle;"><tr><td></td><td></td><td></td><td></td></tr><tr><td></td><td></td><td></td><td></td></tr></table> Duration [Months] <table border="1" style="display: inline-table; vertical-align: middle;"><tr><td></td><td></td><td></td><td></td></tr><tr><td></td><td></td><td></td><td></td></tr></table>																
1	d	d	m	m	1	9	5	7	Months	0	6	Went back to work																	
									Years	0	1																		

In this example this woman can only recall the year of birth of her first child, whom she breast fed for 18 months [1 year and 6 months, a year and a half]. Her reason for stopping was that she went back to work.

Figure 3. Example from section 3H

Child	Date of birth of child								Duration Breastfeeding			Reason for stopping	YoB <table border="1" style="display: inline-table; vertical-align: middle;"><tr><td></td><td></td><td></td><td></td></tr><tr><td></td><td></td><td></td><td></td></tr></table> Duration [Months] <table border="1" style="display: inline-table; vertical-align: middle;"><tr><td></td><td></td><td></td><td></td></tr><tr><td></td><td></td><td></td><td></td></tr></table>																
4	0	2	0	1	1	9	6	3	Months	0	2	Child died																	
									Years	0	0																		

In this example this woman has provided a complete date of birth for her fourth child. She only breastfed this child for two months, she stopped because the child passed.

For the questions about a woman's age at first birth, last birth and the age she stopped breastfeeding please make a note as we can calculate her age at these time points using her date of birth and the birth dates of her children. For the age when she stopped breastfeeding don't forget to add the length of time she breastfed. Some women will have stopped breastfeeding before they have completed their families. For example if woman last breastfed with her third child, but bottle fed children 4 and 5 her age when she stopped breastfeeding is her age when she delivered child number 3 plus the length of time she breastfed that child.

Section 4 – Use of hormonal contraceptives or HRT

This section asks women about their use of hormonal contraceptives either oral (like the pill) or injectable (like depo or petogen). It also asks about the use of Hormone Replacement Therapy (HRT). HRT is a type of medication some women may have been given during their menopause (“change of life”). If a woman doesn’t know what HRT is she is unlikely to be using it. Remember to record 9 for any answer where a woman is unsure or doesn’t know.

Section 5 – Additional Breastfeeding Information

It is possible that some women may have breastfed children that were not their own. This may be because a woman donated breast milk or perhaps took on the responsibility of breastfeeding another child (one she did not give birth to). This might change the total amount of time a woman would have breastfed. In section 2 we ask women if they have ever breastfed a child that was not their own (this should include women who donated breast milk). If they respond “YES” [recorded as 1 in question 3 G] please remember to complete this section. Interviewers should ask for assistance from Debbie or Serena if they struggle with this section.

FAQ

What should I do if I see a woman is 39 years old?

If she will turn 40 in the next 6 months she should be given a questionnaire

What do I do if a woman can’t remember a date or other bit of information?

Record the number 9 in the appropriate block

What do I do if I see a woman who is over 55 and is still having her periods or who says she stopped having periods but now they’ve come back?

Please ask for assistance and review from Serena or Debbie

A woman tells me she was pregnant 7 times but only has 2 children, what do I put down?

Please ask her for more information about her pregnancies. If she says she was pregnant 7 times but lost 5 of the pregnancies before 6 months then please record 7 for “How many times have you been pregnant?” and 2 for “How many children have you delivered?” Please fill in section 3H for the 2 children she gave birth to.

A woman says she had a child but it died at birth, is that a still birth or a live birth?

If the child breathed at birth it’s a live birth, but if the woman is unsure as the child died right around the time it was born you may record it as a stillbirth. Again if unsure ask for assistance

If there are any other questions please ask for assistance

Appendix E

Other work during this thesis

E.1 List of presentations during this thesis

- The long term effects of childbearing and breastfeeding on body mass index in middle aged women results from the Million Women Study . Society for Social Medicine Annual Meeting: Newcastle (Oral), 2009 September

Abstracts

adiposity, were measured directly. Study participants were traced, mailed a health questionnaire in 1962/1966 (mean age 45.1 years) which included enquiries regarding self-reported height and weight, and were followed for subsequent mortality experience – with > 99% completion – until the end of 1998. BMI was categorised into quartiles in the present analyses.

Setting: USA.

Participants: 14 638 men enrolled in Harvard University in the given years, who completed the subsequent health questionnaire, and whose vital status could be ascertained.

Main Outcome Measure: CHD death.

Results: Over a maximum of 82.5 years of follow-up (median 56.5 years), there were 1401 deaths from CHD. Following adjustment for age and other CHD risk factors (cigarette smoking, physical activity, blood pressure) at college entry, relative to the lowest weight quartile (mean BMI = 18.7 kg/m²), there was an elevated risk of CHD mortality in men in the highest quartile (mean BMI = 25.0 kg/m²; hazards ratio 1.28, 95% CI 1.10 to 1.49) but not the intermediate groups. Following additional control for BMI in middle-age, this increased CHD risk in the highest quartile was eliminated (1.03; 0.87 to 1.21).

Conclusion: In this cohort, higher BMI in early adulthood was associated with an elevated risk of CHD mortality several decades later but this effect appeared to be entirely mediated via BMI in middle-age.

056 THE LONG TERM EFFECTS OF CHILDBEARING AND BREASTFEEDING ON BODY MASS INDEX IN MIDDLE AGED WOMEN RESULTS FROM THE MILLION WOMEN STUDY

¹K Bobrow, ²M Quigley, ³J Green, ⁴G Reeves, ⁵V Beral, on behalf of the Million Women Study Collaborators ¹Department of Public Health, University of Oxford, Oxford, UK; ²National Perinatal Epidemiology Unit, University of Oxford, Oxford, UK; ³Cancer Epidemiology Unit, University of Oxford, Oxford, UK

doi:10.1136/jech.2009.096727d

Objectives: To examine the relation between childbearing and breastfeeding and subsequent body mass index (BMI) in middle-aged women.

Design: Cross-sectional analysis within a large prospective study.

Participants: One million UK women recruited into the Million Women Study during 1996–2001. Participants completed questionnaires collecting information on personal and lifestyle characteristics, medical and reproductive history including questions on childbearing and breastfeeding, anthropometry, physical activity, and socio-demographic in factors.

Main Outcome Measures: Mean BMI stratified by parity and breastfeeding and adjusted for age, geographical region, socio-economic status, smoking status, strenuous physical activity, age at first birth, and years since last birth.

Results: 980 474 women were included in the main analysis. 87% of the study population were parous, mean parity was 2.1 (1.2) and 68% of parous women ever breastfed. The mean lifetime duration of breastfeeding per child was 3.8 (SD 3) months. Mean BMI increased significantly with each birth from 25.8 (95% CI 25.7 to 25.8) for nulliparous to 28.1 (95% CI 28.1 to 28.3) for women with five or more births. Parous women who breastfed had significantly lower BMIs than their non-breastfeeding counterparts. The effect was attenuated by adjusting for socio-economic status but remained significant even after full adjustment. The mean BMI decreased as lifetime duration of breastfeeding increased: compared with women who never breastfed, mean BMI decreased by 0.8 (95% CI 0.71 to 0.82) kg/m² in those with a lifetime breastfeeding duration of more than 9 months. This relationship was statistically significant ($p < 0.0001$) and maintained at each parity level.

Conclusions: In this analysis including one million middle-aged women in the UK we found that BMI increased with increasing parity, but that this increase would be offset if women breastfed. These relationships were independent of socioeconomic status,

geographic region, smoking, exercise, age at first birth and time since last birth. These findings contribute to the body of evidence that childbearing and breastfeeding have sustained long term effects on the health status of women.

Ethnicity and young people

057 EXPLAINING THE MENTAL HEALTH ADVANTAGE OF BRITISH INDIAN CHILDREN

A Goodman, V Patel, D Leon. Department of Epidemiology and Population Health, London School of Hygiene and Tropical Medicine, London, UK

doi:10.1136/jech.2009.096727e

Objective: To investigate the causes of the lower rate of mental disorders diagnosed in British Indian children compared to White children.

Design: Secondary analysis of two population-based, cross-sectional surveys of child mental health (the British Child and Adolescent Mental Health Surveys of 1999 and 2004).

Setting: Great Britain (nationally-representative sample). Mental health outcome measures: Parent Strengths and Difficulties Questionnaire (SDQ); teacher SDQ; child SDQ; multi-informant clinical diagnosis.

Participants: 16 449 White and 419 Indian children aged 5–16. The detailed multivariable models focused on the 13 868 White and 361 Indian children from England whose parents had completed an SDQ in English.

Results: There was a substantially lower prevalence ($p < 0.001$) of any mental disorder in Indians (3.4%, 95% CI 1.9 to 5.9) compared to Whites (9.6%, 95% CI 9.1 to 10.2). Strong evidence ($p < 0.002$) of an Indian advantage for externalising problems/disorders was consistently observed for the parent, teacher, child SDQs and for clinical diagnosis. Detailed psychometric analyses of the SDQ and clinical interview measures provided no evidence that this Indian mental health advantage could be explained by a measurement bias in the assessment of mental health. In multivariable analyses the unexplained difference between Indians and Whites decreased somewhat after adjusting for the fact that Indian children were more likely to live in two-parent families (92.2% vs. 65.4%) and less likely to have academic difficulties (e.g. 2.9% vs. 8.6% for parent-reported learning difficulties). In models adjusting for a larger number of child, family, school and area variables the difference reduced only by about a quarter (e.g. from 1.08 to 0.75 SDQ points on the parent SDQ) and remained highly significant ($p < 0.001$). There was little or no evidence of an ethnic difference for internalising problems/disorders in unadjusted or adjusted models.

Conclusions: The mental health difference between Indian and White children is specific to a substantial advantage for externalising disorders, and this advantage appears to be real rather than due to a reporting bias. This advantage is largely unexplained by major risk factors for child mental health problems available in this dataset. Further qualitative and quantitative research into the causes of this advantage has the potential to yield insights which could improve the mental health of children of all ethnic groups.

058 EMERGENCE OF ETHNIC DIFFERENCES IN BLOOD PRESSURE IN ADOLESCENCE: THE DETERMINANTS OF ADOLESCENT SOCIAL WELL-BEING AND HEALTH LONGITUDINAL STUDY

¹E Lenguaerand, ¹S Harding, ¹M Whitrow, ¹M Maynard, ²K Cruickshank, ¹G Der, ¹A Teyhan, ¹Medical Research Council, Social and Public Health Sciences Unit, University of Glasgow, Glasgow, UK; ²Division of Cardiovascular and Endocrine Sciences, University of Manchester, Manchester, UK

doi:10.1136/jech.2009.096727f

Objective: To examine ethnic differences in changes in Blood Pressure (BP) between early and late adolescence in the UK.

- DPhil work-in-progress seminars. The long term effects of childbearing on women's health: Oxford (Oral), 2009 October
- CEU unit seminar. Historical trends in breastfeeding in the UK: Oxford (Oral), 2009 March

E.2 Publications during this thesis

- Persistent effects of childbearing on body mass index in postmenopausal UK women: cross-sectional analysis in a large population - submitted for publication