Body Size, Body Composition, and Physical Activity in Relation to Breast Cancer Risk



Wenji Guo

St Edmund Hall

University of Oxford

A thesis submitted for the degree of

Doctor of Philosophy

Hilary 2018

Acknowledgements

I thank my supervisors, Tim Key and Gill Reeves for their support, encouragement, and understanding throughout what has been an unpredictable but very rewarding journey. I am incredibly lucky to have had such an insightful and engaging supervision team. I also thank Valerie Beral for the opportunity to study at the Cancer Epidemiology Unit.

I would also like to thank my mentors, Donald Chambers and Rhonna Cohen. It has been a privilege to learn from them and get to know them better throughout my Oxford years. Every interaction continues to inspire me toward further scientific pursuits.

I am grateful for my colleagues at the Cancer Epidemiology Unit, who have made this experience such a pleasure. In particular, I thank my extraordinary friend and colleague, Kezia Gaitskell, who has helped me in more ways than I could capture in these acknowledgements.

I have also benefited immensely from discussions and support from Karl Smith Byrne, who made a stressful summer of diving into the deep end of genetic epidemiology ultimately very fun. I would also like to thank Kirstin Pirie, Kathryn Bradbury, Paul Appleby, Ben Cairns, Lucy Wright, Julie Schmidt, Kate Coffey, Georgina Fensom, Aurora Perez-Cornago, Ruth Travis, Ellie Watts, Rob Smith, Jason Lacombe, Angel Wong, Jae Won Suh, Ai Seon Kuan, and Yvonne Walburga.

I am also grateful to the Clarendon Fund and St. Edmund Hall for their financial support and for the opportunity to meet so many friends throughout the years.

Finally, I would like to express my deepest appreciation to my parents and my partner, Tyler Jones, for their love and support.

Abstract

Body Size, Body Composition, and Physical Activity in Relation to Breast Cancer Risk
Wenji Guo, St. Edmund Hall, University of Oxford, DPhil, Hilary Term 2018

Accurately characterising the associations between adiposity, physical activity, and breast cancer risk has been limited by measurement error and confounding. This thesis investigates risk factors for breast cancer, with a focus on body size, body composition, and physical activity, within the UK Biobank, a new prospective cohort with comprehensive assessment of baseline characteristics, direct measures of body fat, and objectively-measured physical activity, which has the potential to address and overcome major limitations of prior studies.

To characterise the patterns of association for breast cancer risk factors in UK Biobank, prospective analyses were conducted in up to 255,083 women. A total of 4,522 women were diagnosed with breast cancer during a mean 5.8 years follow-up. Associations between established risk factors and breast cancer incidence were consistent with prior studies.

All measures of adiposity showed similar cross-sectional associations with self-reported physical activity, which were further investigated using objective, accelerometer measures of physical activity that were available on a subset of 42,992 women in the analyses, 3-5 years after end of recruitment. Compared to the non-linear associations seen with self-reported physical activity, accelerometer-measured physical activity showed nearly twofold larger, inverse, linear dose-response relationships with adiposity.

To investigate whether the effect of genetic susceptibility, as measured by a genetic risk score (GRS) of 70 body mass index (BMI)-associated single nucleotide polymorphisms, on objectively measured BMI was modified by physical activity, cross-sectional analyses of GRS and BMI were performed by physical activity level. Comparing the top to bottom fifth of GRS, the most inactive participants had a 2.23 kg/m^2 higher BMI while the most active participants had only a 1.33 kg/m^2 higher BMI. The results using accelerometer rather than self-reported physical activity provided stronger evidence that genetic predisposition to adiposity can be largely overcome by greater physical activity.

Prospective analyses of various indices of adiposity and breast cancer risk were performed to assess whether any were better predictors of risk, among women who were not current users of hormone replacement therapy. Women in the top compared to bottom quartile of body fat mass were at a 70% increased risk of breast cancer (relative risk (RR) 1.70; 95% confidence interval (CI) 1.52-1.90). All measures of body size and composition showed similar associations with breast cancer risk, suggesting impedance measures of fat were not substantially better indicators of risk than anthropometric measures. After adjusting for body fat mass, central adiposity was no longer associated with breast cancer risk.

Physical activity was inversely associated with risk in both pre- and postmenopausal women. Adjusting for body fat mass slightly strengthened the inverse association in premenopausal women (RR 0.79; 95% CI 0.65-0.95) and attenuated the association in postmenopausal women (RR 0.84; 95% CI 0.72-0.97), per 5 milli-gravity of accelerometer-based physical activity, suggesting that the remaining protective association may be mediated by non-adiposity related factors.

In conclusion, all indices of adiposity were similarly associated with physical activity, and associations were stronger with physical activity measured by accelerometer compared to questionnaire. Physical activity was an effect modifier of genetic susceptibility to adiposity. Adiposity was associated with increased breast cancer risk in postmenopausal women and physical activity was associated with reduced breast cancer risk in both pre- and postmenopausal women. The mechanisms underlying the protective association between physical activity and breast cancer risk are not entirely explained by adiposity and remain unclear.

Contents

1 INTRODUCTION **1**
- 1.1 Rationale . . . 1
- 1.2 Research aims . . . 3
- 1.3 Outline of the thesis . . . 3
- 1.4 Development of thesis and author's role . . . 6
- 1.5 Publications relating to this thesis . . . 7

2 BREAST CANCER EPIDEMIOLOGY **9**
- 2.1 Introduction . . . 9
- 2.2 Geographic trends, temporal trends, and age . . . 10
  - 2.2.1 Geographic trends and migrating populations . . . 10
  - 2.2.2 Temporal trends . . . 10
  - 2.2.3 Age . . . 11
- 2.3 Anthropometric factors . . . 11
  - 2.3.1 Height . . . 11
  - 2.3.2 Body mass index . . . 12
  - 2.3.3 Early-life adiposity . . . 13
  - 2.3.4 Weight change patterns . . . 14
- 2.4 Lifestyle factors . . . 15
  - 2.4.1 Physical activity . . . 15
  - 2.4.2 Diet . . . 15
  - 2.4.3 Alcohol . . . 16
  - 2.4.4 Smoking . . . 16
- 2.5 Reproductive factors . . . 17
  - 2.5.1 Age at menarche . . . 17
  - 2.5.2 Menstrual cycle . . . 18
  - 2.5.3 Childbearing . . . 18
  - 2.5.4 Breastfeeding . . . 19
  - 2.5.5 Oral contraceptive use . . . 20
  - 2.5.6 Age at menopause . . . 20
  - 2.5.7 Hormone replacement therapy . . . 21

2.5.8 Endogenous hormones . . . 22
2.5.9 Hormone receptor status . . . 23
2.6 Clinical factors . . . 24
2.6.1 Mammographic density and benign breast disease . . . 24
2.6.2 Treatment for infertility . . . 25
2.6.3 Ionising radiation . . . 25
2.7 Family history and genetic factors . . . 25
2.7.1 Family history . . . 25
2.7.2 High-penetrance genes . . . 26
2.7.3 Low and moderate-penetrance genes . . . 27
2.7.4 Gene-environment interaction and breast cancer risk . . . 27
2.8 Other factors . . . 28

3 UK BIOBANK AND METHODS 29
3.1 Introduction . . . 29
3.2 Design and methods . . . 29
3.3 Recruitment . . . 31
3.3.1 Data collection . . . 31
3.3.2 Baseline questionnaire . . . 32
3.3.3 Body size and composition measures . . . 32
3.3.4 Blood sample collection and storage . . . 33
3.3.5 Genetic data . . . 33
3.4 Ethics . . . 34
3.5 Follow-up . . . 34
3.5.1 Outcome of incident breast cancer . . . 35
3.5.2 Repeat assessments and additional measures at re-assessment 35
3.6 Statistical analyses . . . 36

4 RISK FACTORS FOR BREAST CANCER IN UK BIOBANK 38
4.1 Introduction . . . 38
4.2 Methods . . . 38
4.2.1 Study design and participants . . . 39
4.2.2 Exposure assessment . . . 40
4.2.3 Ascertainment of cancer cases . . . 42
4.2.4 Statistical analysis . . . 42
4.3 Results . . . 44
4.3.1 Anthropometric factors . . . 44
4.3.2 Lifestyle factors . . . 45
4.3.3 Reproductive factors . . . 45
4.3.4 Other factors . . . 46
4.4 Discussion . . . 46

4.5 Conclusion 50

5 SELF-REPORTED PHYSICAL ACTIVITY IN RELATION TO BODY SIZE AND COMPOSITION 60
5.1 Introduction 60
5.2 Methods 61
5.2.1 Data source 61
5.2.2 Study participants 62
5.2.3 Physical activity 62
5.2.4 Anthropometry and body composition 64
5.2.5 Statistical analyses 64
5.3 Results 66
5.4 Discussion 68
5.5 Conclusion 72

6 ACCELEROMETER AND QUESTIONNAIRE MEASURES OF PHYSICAL ACTIVITY IN RELATION TO BODY SIZE AND COMPOSITION 83
6.1 Introduction 83
6.2 Methods 84
6.2.1 Data source 84
6.2.2 Study participants 84
6.2.3 Self-reported physical activity 85
6.2.4 Accelerometer-measured physical activity 86
6.2.5 Anthropometry and body composition 87
6.2.6 Statistical analyses 87
6.3 Results 89
6.4 Discussion 91
6.5 Conclusion 94

7 GENE BY ENVIRONMENT INTERACTIONS IN ADIPOSITY 103
7.1 Introduction 103
7.2 Methods 104
7.2.1 Data source 104
7.2.2 Study sample 104
7.2.3 Assessment of BMI, physical activity, and screen time behaviour 105
7.2.4 Genotyping 106
7.2.5 Genetic risk score (GRS) computation 106
7.2.6 Statistical analyses 107
7.3 Results 108

7.3.1 Characteristics of study participants and association of weighted genetic risk score with BMI . . . 108
7.3.2 Adjusted geometric means of BMI by genetic risk score and total physical activity . . . 108
7.3.3 Association of genetic risk score with BMI and odds of being obese, by physical activity level . . . 109
7.3.4 Associations of physical activity with BMI by genetic risk score: comparison of self-reported versus accelerometer data 110
7.3.5 Association of genetic risk score with BMI and odds of being obese, by leisure screen time . . . 111
7.4 Discussion . . . 111
7.5 Conclusion . . . 115

8 INDICES OF ADIPOSITY AND BREAST CANCER RISK IN POSTMENOPAUSAL WOMEN 130
8.1 Introduction . . . 130
8.2 Methods . . . 131
8.2.1 Data source . . . 131
8.2.2 Study participants . . . 132
8.2.3 Anthropometry and body composition . . . 132
8.2.4 Ascertainment of cancer cases . . . 133
8.2.5 Statistical analyses . . . 134
8.3 Results . . . 137
8.4 Discussion . . . 139
8.5 Conclusion . . . 142

9 PHYSICAL ACTIVITY AND BREAST CANCER RISK 153
9.1 Introduction . . . 153
9.2 Methods . . . 154
9.2.1 Data source . . . 154
9.2.2 Study participants . . . 154
9.2.3 Physical activity . . . 155
9.2.4 Anthropometry and body composition . . . 156
9.2.5 Ascertainment of cancer cases . . . 156
9.2.6 Statistical analyses . . . 156
9.3 Results . . . 159
9.4 Discussion . . . 162
9.5 Conclusion . . . 165

10 DISCUSSION **174**
10.1 Aims of thesis . . . 174
10.2 Findings in context . . . 175
10.2.1 Risk factors for breast cancer in UK Biobank . . . 175
10.2.2 Self-reported physical activity in relation to body size and composition . . . 177
10.2.3 Accelerometer and questionnaire measures of physical activity in relation to body size and composition . . . 178
10.2.4 Gene by environment interactions in adiposity . . . 180
10.2.5 Indices of adiposity and breast cancer risk in postmenopausal women . . . 181
10.2.6 Physical activity and breast cancer risk . . . 182
10.3 Methodological considerations . . . 184
10.3.1 Measurement error . . . 185
10.3.2 Confounding . . . 186
10.3.3 Reverse causality . . . 187
10.3.4 Generalisability and representativeness . . . 187
10.4 Public health implications . . . 188
10.5 Recommendations for further research . . . 190
10.6 Conclusion . . . 192

A QUESTIONS FROM UK BIOBANK TOUCHSCREEN QUESTIONNAIRE RELEVANT TO THIS THESIS **193**

B INTERNATIONAL PHYSICAL ACTIVITY QUESTIONNAIRE SCORING GUIDELINES **199**

C PAPERS PUBLISHED FROM DPHIL WORK **203**

Bibliography **229**

List of Tables

4.1 Baseline characteristics of all women included in the analysis and of women who developed breast cancer 51
4.2 Relative risks (95% CI) for breast cancer in relation to sociodemographic factors at baseline 53
4.3 Relative risks (95% CI) for breast cancer in relation to anthropometric and lifestyle factors at baseline 55
4.4 Relative risks (95% CI) for breast cancer in relation to reproductive factors and other factors at baseline 57

5.1 Summary characteristics for the 38,201 premenopausal women participants of UK Biobank by physical activity level 73
5.2 Summary characteristics for the 94,592 postmenopausal women participants of UK Biobank by physical activity level 74
5.3 Characteristics of manual and non-manual workers for pre- and postmenopausal women 75
5.4 Age-adjusted Pearson's partial correlation coefficients for the 38,201 premenopausal women participants of UK Biobank 76
5.5 Age-adjusted Pearson's partial correlation coefficients for the 94,592 postmenopausal women participants of UK Biobank 77

6.1 Baseline characteristics according to fifths of accelerometer-measured physical activity 95
6.2 Spearman correlation between self-reported physical activity (in MET-hours) and accelerometer-measured physical activity (in milligravity), according to participant characteristics 96
6.3 Association of self-reported questionnaire and accelerometer-measured physical activity with measures of adiposity 97
6.4 Association of accelerometer-measured physical activity with body fat percentage, measured by bioelectrical impedance and dual energy x-ray absorptiometry (DXA) 98

7.1 Characteristics of the 70 single nucleotide polymorphisms (SNPs) associated with body mass index at genome wide significance, included in the analyses and the 6 that were excluded 117
7.2 Characteristics of study participants 120
7.3 Association of fifths of genetic risk score with body mass index (kg/m^2) 121
7.4 Adjusted geometric means and (95% confidence intervals) of body mass index (kg/m^2) by fifths of genetic risk score and physical activity levels 122
7.5 Association of fifths of genetic risk score with body mass index (kg/m^2), by physical activity level 123
7.6 Association of fifths of physical activity with body mass index (kg/m^2), by genetic risk score 124
7.7 Association of fifths of genetic risk score with body mass index (kg/m^2), by leisure screen time 125

8.1 Baseline characteristics for the 162,691 postmenopausal participants of UK Biobank by various measures of adiposity 143
8.2 Age-adjusted Pearson's partial correlation coefficients for the 162,691 postmenopausal participants 144
8.3 Association of height and overall adiposity measures with invasive breast cancer risk among never and former postmenopausal hormone replacement therapy users 145
8.4 Association of central adiposity measures and invasive breast cancer risk among never and former postmenopausal HRT users 147
8.5 Linear trend for adiposity and breast cancer risk using repeat assessment of baseline measures by impedance compared to DXA measures 149
8.6 Association of body size/composition indices with invasive breast cancer risk in women who have been postmenopausal for less than 12 years versus 12 or more years 150

9.1 Characteristics of the study population at recruitment, according to self-reported physical activity and menopausal status 166
9.2 Association between self-reported and accelerometer-based physical activity and risk of invasive breast cancer by menopausal status . . 167
9.3 Association between self-reported and accelerometer-based physical activity and risk of invasive breast cancer by body mass index . . . 168
9.4 Association between self-reported physical activity and risk of invasive breast cancer, excluding the first two years of follow-up 169

List of Figures

4.1 Participants inclusion/exclusion flow diagram for the analyses of risk factors for breast cancer in UK Biobank 59

5.1 Participant inclusion/exclusion flow diagram for analyses of physical activity in relation to body size and composition 78
5.2 Distribution of physical activity (excess MET-h/week) in premenopausal (A) and postmenopausal (B) women. 79
5.3 Association between physical activity and body size/composition by menopausal status for the 38,201 premenopausal and 94,592 postmenopausal women of UK Biobank 80
5.4 Association between physical activity and A) BMI and B) Body fat percentage by menopausal status, separately in manual workers and non-manual workers. 81
5.5 Association between total physical activity and body size/composition by categories of vigorous physical activity for 38,201 premenopausal and 94,592 postmenopausal women. 82

6.1 Participant inclusion/exclusion flow diagram for analyses of accelerometer-measured and self-reported physical activity in relation to adiposity 99
6.2 Association of accelerometer-measured and self-reported physical activity with adiposity variables 100
6.3 Association of accelerometer-measured physical activity with body fat percentage measured by impedance and dual-energy X-ray-absorptiometry (DXA), 1,272 women included in analysis 101
6.4 Association of self-reported and accelerometer-measured physical activity with adiposity variables by age group 102

7.1 Participant inclusion/exclusion flow diagram for analyses of gene-environment interactions in adiposity 126
7.2 Association of fifths of genetic risk score (based on quintile cut-points) and body mass index (BMI), by physical activity level in metabolic equivalent hours per week (MET-hrs/wk) 127

7.3 Association of physical activity, measured by A) self-reported questionnaire and B) accelerometer, with body mass index (BMI) by thirds of genetic risk score (GRS) 128
7.4 Association of fifths of genetic risk score and body mass index (BMI), by screen time behaviour 129

8.1 Participant inclusion/exclusion flow diagram for analyses of adiposity and breast cancer risk in postmenopausal women 151
8.2 Association between various measures of adiposity and postmenopausal breast cancer risk: (A) Body mass index, (B) Body fat mass, (C) Body fat percentage, (D) Waist circumference, (E) Hip circumference, (F) Waist-to-hip ratio 152

9.1 Participant inclusion/exclusion flow diagram for analyses of physical activity in relation to breast cancer risk 170
9.2 Association between accelerometer-based physical activity and risk of invasive breast cancer 171
9.3 Accelerometer-based physical activity median within each category of questionnaire-based physical activity, according to BMI (in units of kg/m^2) group 172
9.4 Relative risk of accelerometer-based physical activity and breast cancer risk by BMI, ever use of HRT, number of births and alcohol intake 173

Chapter 1

INTRODUCTION

1.1 Rationale

Worldwide, breast cancer is the most common cancer, with approximately 2.4 million new cases diagnosed annually. The lifetime odds of developing breast cancer are 1 in 9 for women in countries with high sociodemographic index[1]. From a public health perspective, although several factors have been linked to breast cancer[2], body size/composition and physical activity are of particular interest since they are among the most modifiable. Based on currently available estimates, 8.7% of breast cancers in United Kingdom (UK) women are due to excess weight and 3.4% due to lack of physical activity[3].

However, accurately characterising the associations between adiposity, physical activity, and breast cancer risk has been hampered by issues such as confounding and measurement error. For example, most large prospective cohort studies have used self-reported anthropometric measures such as body mass index (BMI) which does not distinguish between lean and fat mass and does not capture changes in

adiposity with age[4,5]. Furthermore, prior studies disagree on whether the distribution of adipose tissue influences breast cancer risk[6]. Physical activity is believed to be important for maintaining and preventing increases in body size and composition[7], yet results from randomised controlled trials have been inconsistent[8] and most large-scale observational studies are based on self-reported physical activity from questionnaire responses, which is prone to both inaccurate recall and reporting bias[9]. Although various potential biological pathways have been proposed[10], physical activity's likely protective association with breast cancer risk and the extent to which the association may be mediated by adiposity remain unclear.

There is thus a need to develop epidemiological analyses of breast cancer risk with improved assessment of exposures. Given the importance of adiposity as a risk factor for breast cancer in postmenopausal women and the proposed link to adipose tissue as the major site of excess oestrogen synthesis in postmenopausal women[11], more detailed and objective measures of fat mass, preferably using repeated assessments, are needed to better characterise the association. The determinants of adiposity itself are complex and it is important to better understand the interplay between genetic and behavioural factors such as physical activity in influencing body size and composition.

Large prospective cohort studies with objective measures of physical activity are needed to reduce the effect of measurement error from self-reported estimates, resolve inconsistencies in the literature regarding the strength of association with breast cancer risk, and clarify whether the protective association is of greater magnitude in different subgroups.

1.2 Research aims

The goal of the work presented in this thesis is to investigate the associations between breast cancer risk and modifiable lifestyle factors with a focus on anthropometric measures of body size, body composition, and physical activity measured by both questionnaire response and accelerometer. I also explore relationships among the risk factors themselves, including whether genetic susceptibility to adiposity is modified by physical activity. I aim to provide novel insights into these associations of public health importance by utilising both direct measures of adiposity and objective measures of physical activity.

1.3 Outline of the thesis

- **Chapter 1** introduces the rationale and research aims of the thesis. It also outlines the specific objectives and contents of each chapter.

- **Chapter 2** provides an overview of breast cancer epidemiology, including both established and putative risk factors. These include age, height, body mass index, early-life adiposity, weight change patterns, physical activity, diet, alcohol, smoking, parity, age at menarche, menstrual cycle, childbearing, age at first birth, breastfeeding, oral contraceptive use, age at menopause, hormone replacement therapy (HRT) use, endogenous hormones, hormone receptor status, mammographic density, benign breast disease, treatment for infertility, ionising radiation, and genetic factors.

- **Chapter 3** provides a description of the design and methods of the UK Biobank, the data source used throughout this thesis.

- **Chapter 4** reports the first results in UK Biobank on the associations between breast cancer incidence and established risk factors, in more than a quarter million UK Biobank women and discusses these findings in relation to the literature.

- **Chapter 5** characterises the cross-sectional associations between self-reported physical activity and direct measures of both total and central adiposity separately in pre- and postmenopausal women using multivariable linear regression. Measures of adiposity include body mass index (BMI), body fat percentage, body fat mass, trunk fat percentage, trunk fat mass, waist circumference, hip circumference, and waist-to-hip ratio (WHR). I also evaluated whether the association between total physical activity and the measures of adiposity varied depending on the proportion of vigorous-to-total physical activity. The work included in this chapter was published in Annals of Epidemiology (June 2015, Volume 25, Issue 6, pages 406-413).

- **Chapter 6** compares the cross-sectional associations of self-reported questionnaire vs. accelerometer-measured physical activity in relation to various measures of adiposity including BMI, waist circumference, and body fat percentage measured by both bioelectrical impedance and dual energy X-ray absorptiometry. Accelerometer-measured physical activity was measured from 2013-2015 on a subset of participants. This chapter also explores how these associations may vary by sociodemographic characteristics such as age and education as well as variations in the strength of correlation between self-reported and accelerometer-measured physical activity according to various sociodemographic characteristics. A manuscript relating to this work has

been prepared for publication.

- **Chapter 7** investigates whether genetic susceptibility to adiposity is modified by physical activity and leisure screen time behaviours. A weighted genetic risk score of 70 single nucleotide polymorphisms (SNPs) that showed genome-wide significant association with BMI was calculated as a measure of genetic susceptibility to adiposity. Furthermore, the associations between physical activity and BMI by genetic risk score were compared using accelerometer versus self-reported physical activity data. This study is a cross-sectional analysis in the UK Biobank cohort of 46,785 women.

- **Chapter 8** investigates the associations between body size, body composition, and breast cancer risk in postmenopausal women. Given that most large prospective cohort studies have used BMI as the measure of adiposity, which is limited by its inability to directly distinguish between fat and lean mass[12], this chapter assesses whether direct measures of fat mass are better indicators of risk than anthropometric measures. A secondary aim of the chapter is to investigate whether central adiposity contributes to breast cancer risk beyond overall adiposity. This study is a prospective analysis in the UK Biobank cohort of 2,913 cases of breast cancer in postmenopausal women that had accrued by 30 November 2014. Work related to this chapter was published in International Journal of Cancer (doi:10.1002/ijc.31394).

- **Chapter 9** investigates the association between physical activity and breast cancer risk using questionnaire-based, self-reported physical activity in conjunction with accelerometer values, available on a subset of participants. A secondary aim of this chapter was assessing to what extent the association

between physical activity and breast cancer risk may be mediated by adiposity. This study is a prospective analysis in the UK Biobank cohort of 3,189 cases of breast cancer in both premenopausal and postmenopausal women that had accrued by 30 November 2014. A manuscript related to the work in this chapter has been prepared for publication.

- **Chapter 10** discusses the main findings of the thesis, highlights the strengths and limitations of these studies, and provides recommendations for future research.

- **Appendix A** contains parts of the touchscreen questionnaire used in the UK Biobank study that are relevant to the work presented in this thesis.

- **Appendix B** contains part of the scoring guidelines for the International Physical Activity Questionnaire relevant to the work presented in this thesis.

- **Appendix C** contains papers published from DPhil work.

1.4 Development of thesis and author's role

All work presented in this thesis, including literature reviews, statistical analysis and chapter text, is my own unless specifically stated otherwise. All work has been conducted under the supervision of Professor Timothy Key and Professor Gillian Reeves. During the data analysis phase of Chapter 5, I also worked with Dr. Kathryn Bradbury (a nutritional epidemiologist at Cancer Epidemiology Unit [CEU]) to define and create the physical activity variables and received statistical advice from Paul Appleby (a senior statistician at CEU). During the data analysis

phase of Chapter 9, I received statistical advice from Kirstin Pirie (a statistical programmer at CEU). The concept of studying body size, physical activity, and breast cancer risk in UK Biobank was initially developed by Professor Tim Key and the approaches to these analyses were subsequently developed in consultation with my supervisors Professor Tim Key and Professor Gill Reeves. As new data became available, we decided to pursue more detailed analyses using the accelerometer data (Chapter 6) and the genetic data (Chapter 7), adding a greater emphasis to the thesis of exploring the complex associations between physical activity and adiposity. The idea of utilising the genetic data in UK Biobank was initially suggested by and discussed with my transfer assessors, Associate Professor Naomi Allen and Associate Professor Ruth Travis. The genetic analysis was conducted throughout June and July 2016, prior to the publication of UK Biobank studies on this topic. I received assistance on the initial cleaning of the genetic data from Dr. Karl Smith Byrne (a former DPhil student at CEU) and further advice from Dr. Ben Cairns (a senior statistical epidemiologist at CEU).

1.5 Publications relating to this thesis

At the time of thesis writing, two manuscripts have been published from work described in this thesis. These are "Physical activity in relation to body size and composition in women in UK Biobank", and "Adiposity and breast cancer risk in postmenopausal women: results from the UK Biobank prospective cohort" (see Appendix C). My co-authors on these papers provided advice on content and methodology as well as comments on the manuscripts both prior to submission and at revision. A manuscript has also been prepared for publication in relation

to work outlined in Chapter 6 and another manuscript has been prepared for publication in relation to work outlined in Chapter 9.

Chapter 2

BREAST CANCER EPIDEMIOLOGY

2.1 Introduction

What follows is a description of the broad epidemiology of breast cancer. Since breast cancer epidemiology is a vast topic, I have cited data mainly from large studies, such as the Million Women Study, and large pooled analyses, rather than attempt to comprehensively review the literature.

The focus of this thesis is on risk factors for breast cancer in women. Breast cancer does occur in men, with risk factors that differ from those seen among women[13].

2.2 Geographic trends, temporal trends, and age

2.2.1 Geographic trends and migrating populations

There is a two to three-fold difference in the incidence of breast cancer when comparing low-risk regions, such as Africa and Asia, to high-risk regions, such as North America and Northern and Western Europe[14]. Breast cancer incidence increases for migrants who move from a region with low breast cancer incidence, such as Asia, to a region with higher breast cancer incidence, such as the United States. This effect is observed in the descendants of migrants. Within one or two generations, the breast cancer risk of the descendants is also similar to that of the native population[15]. This suggests that the differences in rates are largely due to lifestyle factors rather than differences in genetic susceptibility. For instance, factors such as age at menarche, parity, age at first birth and subsequent births, and adult weight may underlie differences in incidence between low-risk and high-risk regions[16].

2.2.2 Temporal trends

Between 1980 and the late 1990s, age-adjusted breast cancer incidence rates increased by approximately 30% in Western countries, mainly as a result of increased use of hormone replacement therapy, changes in reproductive patterns, and increased screening. After reports of increased breast cancer risk among users of hormone replacement therapy in the early 2000s, the increase in breast cancer incidence slowed. Plateaus in screening participation may also be a factor in the stabilisation of incidence rates[14].

Breast cancer incidence rates in lower and middle-income countries continue to rise, likely due to a combination of changing reproductive patterns, increasing obesity, decreasing physical activity, as well as increased awareness and screening[14,17].

2.2.3 Age

Breast cancer incidence increases more rapidly with age during the reproductive years. After around age 50, which is the average age at menopause, incidence increases at a comparatively slower rate[2].

2.3 Anthropometric factors

2.3.1 Height

Height is positively associated with increased risk of breast cancer. A recent study containing a meta-analysis of prospective studies found a stronger association among studies with measured height compared to studies with self-reported height (relative risk (RR) 1.20, 95% confidence interval (CI) 1.17-1.23 for studies with measured height and RR 1.17, 95% CI 1.15-1.19 for studies with self-reported height; P for heterogeneity = 0.001) per 10 cm increase[18]. Within the same study, the Mendelian randomisation approach was also used and relative risk per 10 cm increase in height was 1.22 (95% CI 1.13-1.32) based on the instrumental variable analysis in all women in the Breast Cancer Association Consortium. When stratified by oestrogen receptor status, the association was seen only in oestrogen receptor positive cancers (P for interaction = 0.02)[18].

Furthermore, adult height may be associated with higher circulating concen-

trations of growth factors such as insulin-like growth factor 1 (IGF-1) during childhood. In childhood and adolescence, circulating concentrations of IGF-1 are strongly positively associated with skeletal growth. Circulating levels of IGF-1 have been observed to be greater among taller compared to shorter women and are associated with increased breast cancer risk[18–20].

2.3.2 Body mass index

In the Million Women Study, for a 10 unit increase in BMI, the relative risk of breast cancer in postmenopausal never users of hormone replacement therapy was 1.40 (95% CI 1.31-1.49)[21]. For postmenopausal women, this increase in breast cancer risk seen with increasing BMI is largely due to the corresponding increase in circulating oestrogens. Since ovarian production of oestrogens stops after menopause, adipose tissue becomes the main source of oestrogen. Adipose tissue is a site of aromatase activity, which catalyses the major step of the pathway for the conversion of androgens into oestradiol. This can result in an approximately 2-fold higher serum oestradiol concentration in obese postmenopausal women when compared to thin postmenopausal women, as observed in a pooled analysis of prospective studies[22,23]. The increased risk of breast cancer with increasing BMI was reduced after adjusting for serum oestrogens. However, the association between BMI and breast cancer risk was not significantly affected by adjustment for androgens[22]. In addition, the association of greater adiposity with increased breast cancer risk is obscured in current users of hormone replacement therapy, likely since circulating oestradiol concentrations would be high in current users regardless of body size[24,25].

There has been considerable evidence in Western populations of a negative association between BMI and breast cancer risk for premenopausal women[26,27]. In premenopausal women, obesity may cause anovulation, leading to decreased oestrogen and progesterone production and cumulative exposure, although other mechanisms may be involved since adjustment for menstrual cycle patterns, infertility due to an ovulatory disorder, likely polycystic ovarian syndrome, and oral contraceptive use did not explain the inverse association between BMI and breast cancer risk in premenopausal women[26]. Another proposed mechanism is that the larger reservoir of adipose tissue in obese women may sequester oestradiol, leading to a fall in blood levels of oestradiol[28].

For postmenopausal breast cancer, ER^+PR^+ tumours are associated with a greater magnitude increased risk with increasing BMI, compared to ER^- PR^-, ER^+ PR^-, and triple negative tumours[29–32]. For premenopausal women, higher adiposity has been associated with reduced risk of ER^+PR^+ breast cancer[32]. Additionally, the Nurses' Health Study II reported that greater adiposity was associated with increased risk of ER-PR- and triple negative breast cancer while the European Prospective Investigation into Cancer and Nutrition found no association[31,33].

2.3.3 Early-life adiposity

Regardless of menopausal status, greater adiposity at younger ages is inversely associated with breast cancer risk for both premenopausal and postmenopausal women, independent of current adult BMI[34]. A possible mechanism underlying this protective association is favourable mammographic density[35]. Other mechanisms that have been suggested include the possibility that higher oestrogen

exposure in pre-pubertal girls[36] may induce breast tissue differentiation, which would eliminate the potential for malignant transformation, and the upregulation of tumour suppressor gene BRCA1 which participates in DNA damage repair[37].

Mendelian randomisation studies support an inverse association between BMI predicted using a genetic risk score and breast cancer risk in both pre- and post-menopausal women[38,39]. The genetic risk score used in these studies may be more predictive of BMI in early-life rather than later adulthood[40].

2.3.4 Weight change patterns

Some studies report that adult weight change patterns may be more informative for cancer risk than solely baseline BMI measured in cohorts of middle-aged adults. The Women's Health Initiative, with 20 years of follow-up, showed an association between weight gain and increased risk of breast cancer in postmenopausal women, independent of baseline BMI[41]. A meta-analysis of prospective observational studies found that each 5 kg increase in adult weight gain was associated with an 11% increased risk of breast cancer among postmenopausal women who were not using hormone replacement therapy[42]. Analyses from the Nurses' Health Study estimated that approximately 17% of ER^+PR^+ breast cancer is attributable to weight gain after age 18 of more than 5 kg[43]. However, it is unclear to what extent weight gain is the driving factor in the association between adiposity and increased breast cancer risk or whether these findings are reflected in the previously discussed protective association of adiposity in early-life and increased risk with adiposity in adult life.

2.4 Lifestyle factors

2.4.1 Physical activity

Overall, several studies suggest an association between physical activity and reduced risk of breast cancer, although these results are based on self-reported physical activity[44]. The exact mechanisms are uncertain, but physical activity may influence endogenous hormone concentrations. Physical activity may be associated with lower circulating levels of oestradiol in premenopausal women[45]. In postmenopausal women, an increase in physical activity has been associated with a decrease in oestradiol and an increase in sex hormone-binding globulin, which results in a lower amount of circulating free oestradiol[46].

Other mechanisms that have been proposed as potential explanations for the association between physical activity and reduced breast cancer risk include improved insulin sensitivity, enhanced immune function, and reduced chronic inflammation[10,47,48].

2.4.2 Diet

While various dietary factors have suspected links to breast cancer, there was only convincing evidence for alcohol intake[49], discussed separately below. Some studies have suggested protective associations for fruit and/or vegetable intake[50,51] and fibre intake[52–54] with breast cancer risk. Analyses of red and processed meat intake in relation to breast cancer have also yielded conflicting results, with one recent study concluding that red but not processed meat was associated with increased breast cancer risk[55] while a recent study, containing both an analysis

in UK Biobank and a meta-analysis of published cohort studies, reported that processed but not red meat was associated with increased breast cancer risk[56]. Results from a recent meta-analysis suggest no association between dietary total fat and fatty acids with breast cancer risk[57].

2.4.3 Alcohol

A recent pooled analysis of 20 prospective studies found a positive association between alcohol consumption and breast cancer risk. Compared to non-drinkers, women who consumed 5-15 grams per day had a 10% increased risk (RR 1.10; 95% CI 1.06-1.14) and women who consumed 30 or more grams per day had a 32% increased risk (RR 1.32; 95% CI 1.23-1.41)[58]. The associations did not differ significantly by ER or PR status, although an earlier meta-analysis that included both cohort studies (all included in the more recent pooled analysis by Jung et al.) and case-control studies reported a stronger association for ER^+ compared with ER^- breast cancer[59].

A number of mechanisms have been proposed to explain alcohol's association with increased breast cancer risk, including increased levels of circulating sex hormones[60]. Animal experiments have demonstrated the toxic and carcinogenic properties of acetaldehyde. Alcohol is metabolised to acetaldehyde, which may cause DNA damage[61,62].

2.4.4 Smoking

In 2012, the International Agency for Research on Cancer concluded that there are several plausible biological mechanisms by which smoking may increase breast

cancer risk, although there was a lack of clear association from epidemiological studies up to that point[63]. Recent epidemiological studies provide support for a modest association between active smoking and breast cancer risk, particularly for women who begin smoking prior to first birth[64–66]. Elevated risk is more strongly correlated with heavy smokers who began smoking earlier in life for a relatively long duration. In a large pooled analysis, the associations did not vary by menopausal status but were stronger for risk of ER^+ compared to ER^- breast cancer[65].

Tobacco smoke contains carcinogens which can enter mammary tissue and may be stored, metabolised, and activated in breast adipose tissue. However, smoking may also exert an anti-estrogenic effect, as reflected by early age at menopause and higher risk of osteoporosis among smokers, which may reduce risk[66].

2.5 Reproductive factors

2.5.1 Age at menarche

In a large pooled analysis, there was a 5% increased risk of breast cancer for every year younger at menarche (RR 1.05; 95% CI 1.04-1.06)[67]. A recent Mendelian randomisation analysis suggests breast cancer risk is inversely and causally related to age at menarche, independent of BMI[68]. Greater childhood body size has a well-established association with earlier age at menarche[69], although childhood adiposity has a protective association with breast cancer risk[34].

Oestrogen, progesterone, and prolactin levels increase markedly around menarche. Early menarche increases lifetime exposure to endogenous sex hormones[70], which could explain an increase in breast cancer risk.

2.5.2 Menstrual cycle

Cycle length and regularity have shown no consistent relationship with breast cancer risk. However, the risk of breast cancer has been hypothesised to increase as a result of more exposure to the luteal phase of the menstrual cycle. Oestrogen and progesterone levels along with breast tissue mitotic activity are highest during the luteal phase. Since the luteal phase has a more constant length than the follicular phase, women with shorter cycles may spend more time in the luteal phase throughout their reproductive lifespan. Furthermore, reduced breast cancer risk has been seen in women with a history of ovulatory infertility, possibly due to less exposure to the luteal phase of the menstrual cycle[71].

2.5.3 Childbearing

Childbearing increases short-term risk but is protective in the long term[72]. To try to disentangle all the various temporal factors, which are interrelated, is difficult. There is an approximately 25% reduction in risk for women with at least one full-term pregnancy compared to nulliparous women and an increased protective effect with more full-term pregnancies[2]. Younger age at first full-term pregnancy is protective possibly because it shortens the interval between menarche and first birth when the breast is highly susceptible to mutations. A further risk reduction is seen with early age at subsequent pregnancies. Late age at first full-term pregnancy and nulliparity are associated with higher risk of breast cancer. Some studies show that women with first birth after age 35 have a higher risk of breast cancer than nulliparous women[73], although this observation may be due to the births in the older women being more recent.

The protective effect of full-term pregnancy may be partially explained by decreased proliferative activity of fully differentiated breast tissue. In animal models, pregnancy and lactation induce complete differentiation of mammary gland epithelial cells, which results in increased capability for DNA repair of the mammary epithelium and activation of genes controlling programmed cell death[74]. Protection conferred by pregnancy has been demonstrated in rat and mouse models. After treatment with a hormone regimen (oestrogen and progesterone, oestrogen alone, or human chorionic gonadotropin), incidence of mammary carcinomas decreases[75]. After treatment with levels of oestradiol and progesterone that mimic pregnancy, four different strains of nulliparous rats were significantly protected against mammary tumorigenesis. Furthermore, microarray expression profiling of mammary tissue from the four rat strains treated with pregnancy-level hormones revealed a gene expression signature associated with reduced breast cancer susceptibility observed in parous but not nulliparous rats[76]. In humans, full-term pregnancy has been demonstrated to induce a specific gene expression signature in the breast that was not observed in women with breast cancer nor in nulliparous women without breast cancer[77].

2.5.4 Breastfeeding

For every 12 months of breastfeeding, the relative risk of breast cancer decreased by 4.3% (95% CI 2.9-5.8) in a pooled reanalysis of 47 studies including women with diverse reproductive and breastfeeding characteristics[78]. It has been suggested that the protective effect of breastfeeding may be stronger in younger women[2]. Lower risk of breast cancer with longer duration of breastfeeding may be partially

attributed to a reduction in cumulative ovarian hormone exposure due to reduced number of ovulatory menstrual cycles[2].

2.5.5 Oral contraceptive use

Relative risk of breast cancer diagnosis increases for women who are currently taking combined oral contraceptives (1.24, 95% CI 1.15-1.33) but decreases after cessation. Ten or more years after cessation, there was no significant excess risk of breast cancer diagnosis (RR 1.01, 95% CI 0.96-1.05). Combined oral contraceptive use was associated with more localized cancers than metastasized cancers[79]. Similarly, a recent large, nationwide prospective study found that compared with never users of oral contraceptives, current and recent users had a 20% increased risk of breast cancer (RR 1.20, 95% CI 1.14-1.26)[80].

2.5.6 Age at menopause

Breast cancer risk increases by 3% for each year older at menopause[67]. Recent Mendelian randomisation analyses confirm an increase in risk for each year older at menopause[81]. Results from a large pooled analysis suggest a rapid decrease in breast cancer risk soon after menopause in women of the same ages[67]. Oestrogen, progesterone, and prolactin levels rapidly decrease around the time of menopause as cycle ovarian hormone production ceases, after which breast epithelium exhibits lower mitotic activity[70]. Adiposity has been observed to modify the association between menopause and breast cancer risk in that the relative risk of breast cancer decreases more quickly after menopause in leaner women, which is likely linked to comparatively lower oestradiol concentrations among women with lower BMI[67].

2.5.7 Hormone replacement therapy

Breast cancer risk increases with increasing duration of use of hormone replacement therapy after menopause. Relative risk of breast cancer diagnosis for current or recent users with duration of use of five or more years was 1.35 (95% CI 1.21-1.49) in a large pooled analysis[82]. This excess risk associated with hormone replacement therapy use decreases after the cessation of use and largely disappears after approximately five years. The effect of hormone replacement therapy on breast cancer risk is greater in lean than in overweight and obese women[82] likely since circulating oestradiol concentrations are already increased in women with high BMI[22]. The Women's Health Initiative (WHI) randomised controlled clinical trial also reported a statistically significant increase in breast cancer incidence and mortality for postmenopausal women who were receiving oestrogen plus progestin (RR 1.25 95% CI 1.07-1.46) compared to placebo[83].

In the Million Women Study, current use was associated with a 66% increased risk of breast cancer compared to never use of hormone replacement therapy at baseline (RR 1.66, 95% CI 1.60-1.72). Use of preparations containing combined oestrogen-progesterone was associated with a greater increased risk (RR 2.00, 95% CI 1.88-2.12) than use of preparations containing oestrogen alone (RR 1.30, 95% CI 1.21-1.40)[84]. However, the WHI reported a non-significant decrease in breast cancer incidence (RR 0.77, 95% CI 0.59-1.01) with use of oestrogen alone in postmenopausal women who had previously undergone hysterectomy[85]; it is not clear why these results are suggestive of a protective effect of oestrogen alone treatment on breast cancer risk. Given that the excess risk of breast cancer associated with oestrogen alone hormone therapy has been observed to be lower with increasing

BMI[82], the discrepancy with prior results from observational studies, such as the Million Women Study, may be explained by a much higher prevalence of obesity in the WHI trial compared to the Million Women Study (45% in the WHI compared to 18% in the Million Women Study[86]).

A more recent study suggests that the magnitudes of association between hormone replacement therapy use and breast cancer risk in most of the published literature are likely to be underestimated due to lack of updated information on hormone replacement therapy use throughout follow-up in prospective studies and the inclusion of women with uncertain age at menopause[87].

2.5.8 Endogenous hormones

Higher levels of endogenous oestrogens and androgens as well as lower levels of sex-hormone binding globulin (SHBG), which result in an increase of bioavailable oestradiol, are associated with increased risk of postmenopausal breast cancer. On average, among six prospective studies, postmenopausal women who developed breast cancer had a 15% higher level of circulating oestradiol compared to postmenopausal women who remained free of breast cancer[88]. The proliferation of breast cells is stimulated directly by oestrogen while androgens may exert an effect through conversion into oestrogens. With the cessation of ovarian synthesis of oestrogens after menopause, the main source of oestrogen for postmenopausal women comes from the adipose tissue where androgens are converted to oestrogens[89].

Increased breast cancer risk has also been associated with higher concentrations of other hormones such as insulin-like growth factor-1 (IGF-1)[2]. IGF-1 is a peptide hormone that stimulates breast epithelial cell growth. Its serum levels are

sensitive to dietary changes[23]. There is some evidence that high circulating IGF-1 is predictive of breast cancer risk in both premenopausal and postmenopausal women. Circulating IGF-1 seems to only be positively associated with oestrogen-receptor-positive tumours[19].

In premenopausal women, risk of breast cancer is associated with increased concentrations of oestradiol, oestrone, calculated free oestradiol, androstenedione, dehydroepiandrosterone sulphate, and testosterone, while luteal phase progesterone and sex-hormone binding globulin show no association[90].

A recent, large consortium performed a nested case-control study of ten prospective cohorts that had collected blood from premenopausal women and found that anti-Mullerian hormone (AMH) was associated with increased breast cancer risk (odds ratio 1.60, 95% CI 1.31-1.94) comparing the top to bottom quartile[91]. AMH is a predictor of age at menopause and its concentration reflects ovarian reserve[92].

Increased concentrations of endogenous hormones may increase breast cancer risk through increasing the mitotic rate of breast epithelial cells, which would increase risk of accumulating mutations. Endogenous hormones may also stimulate growth of early tumours[90].

2.5.9 Hormone receptor status

Breast cancer is a heterogeneous disease characterised by different subtypes based on molecular and genetic characteristics of the tumour cells. Breast cancer tumours differ by their expression of oestrogen receptor, progesterone receptor, and human epidermal growth factor receptor 2. Due to prognostic differences in patient outcomes, tumours are routinely clinically stratified according to the level

of expression of these receptors[93]. Oestrogen and progesterone receptor-positive breast cancers tend to be more responsive to hormonal treatments[94].

Furthermore, hormone receptor status underlies some differences in breast cancer risk factors. Risk factors such as height[18], being nulliparous, delayed childbearing, current use of hormone replacement therapy, and increased age at menopause show stronger and more consistent associations with hormone receptor positive breast cancers[93].

2.6 Clinical factors

2.6.1 Mammographic density and benign breast disease

Mammographic density is strongly associated with increased risk of breast cancer. Increased breast density is associated with a 4.7-fold (95% confidence interval, 3.0-7.4) increased risk of breast cancer for the most dense ($\geq$75%) compared with the least dense ($<$10%) category[95]. Density is defined by the proportion of epithelial and stromal breast tissue in relation to non-dense fatty tissue. Since breast cancers are initiated in epithelial cells, a greater proportion of epithelial tissue may indicate a greater number of cells susceptible to carcinogenesis[96].

Since mammographic density is associated with a number of lifestyle and reproductive factors and with breast cancer risk, it may be an intermediate marker of breast cancer risk through which lifestyle and reproductive factors may influence breast cancer development[97].

Risk also increases substantially with history of proliferative benign breast disease. Women with both proliferative lesions with atypia and high breast density

were at the highest breast cancer risk. However, women with low breast density, even if biopsy results showed presence of proliferative lesions with or without atypia, were at low breast cancer risk[98].

2.6.2 Treatment for infertility

Treatment for infertility may include exposure to various hormonally active drugs. However, there is no clear association between fertility drugs and breast cancer risk[2].

2.6.3 Ionising radiation

Exposure to high-dose ionising radiation increases breast cancer risk in a dose-dependent manner. The highest risk has been observed when exposure occurs before approximately age 20 while none to minimal risk has been reported for women exposed after menopausal ages[99].

2.7 Family history and genetic factors

2.7.1 Family history

Family history is one of the most important risk factors for breast cancer, along with age and gender (incidence is much higher in women than in men[13]). In a large pooled analysis, risk ratios for breast cancer were higher for a greater number of affected first-degree relatives. Compared with no affected first-degree relative, the risk ratio ranged from 1.80 (99% CI 1.69-1.91) for women with one affected first-degree relative to 3.90 (99% CI 2.03-7.49) for women with three affected first

degree relatives[100]. In the Nurses' Health Study, women whose mother or sister had been diagnosed with breast cancer before age 50 had a relative risk of 1.70 (95% CI 1.48-1.95) while women whose mother or sister had been diagnosed after age 50 had a comparatively lower excess risk of breast cancer (RR 1.30, 95% CI 1.27-1.54)[101]. The observed associations between family history in first degree relative and breast cancer risk may be partly explained by mutations in cancer susceptibility genes, discussed below.

2.7.2 High-penetrance genes

Penetrance refers to the proportion of individuals with a specific mutation that will develop the disease. Rare genetic variants for cancer susceptibility generally have higher penetrance (relative risk of carriers versus non-carriers of 5 to greater than 20) than common genetic variants[102]. Individuals with high-penetrance alleles generally have greater than 50% lifetime risk of breast cancer. BRCA1 and BRCA2 are tumour-suppressor genes with important roles in DNA repair. Mutations in BRCA1 and BRCA2 result in high risks for breast, ovarian, and prostate cancer. Deleterious alleles in these two genes explain about 84% of families containing four or more cases of breast cancer diagnosed under the age of 60. Among carriers of BRCA1 mutations who develop breast tumours, more than 90% of the tumours are oestrogen receptor negative[103]. In affected families, early onset breast cancer may be an autosomal dominant condition present throughout multiple generations. Other rare, high penetrance mutations include those in the following genes: TP53 as the cause of Li-Fraumeni syndrome, PTEN as the cause of Cowden syndrome, and LKB1 as the cause of Peutz-Jeghers syndrome[104].

2.7.3 Low and moderate-penetrance genes

The observation that a large proportion of familial, multiple-case clusters of cancers cannot be explained by mutations in high-penetrance genes such as BRCA1 and BRCA2 indicates that susceptibility to common cancers is polygenic in nature. Compared to the high-penetrance variants, low-penetrance variants are more common and are associated with a comparatively lower relative risk (lower than 1.5)[102]. Many low-penetrance genes contribute to increased risk. Low-penetrance alleles indicate an approximately 10-20% lifetime risk and explain about 14% of familial risk of breast cancer. Common, low penetrance alleles tend to be more strongly associated with oestrogen receptor positive breast tumours[103].

Moderate-penetrance alleles are rare and are associated with a twofold to fourfold risk of breast cancer. They confer lifetime risks of 20% or more and explain less than 25% of familial risk of breast cancer[103].

2.7.4 Gene-environment interaction and breast cancer risk

Although it is believed that genetic and environmental factors may interact to modify breast cancer risk, studies to date may still be underpowered and there is no evidence of strong gene-environment interactions[105]. However, interactions of small to moderate magnitude have been reported, such as those between LSP1 and parity as well as CASP8 and alcohol consumption[106]. More recently, a few more weak gene-environment interactions have been reported with current smoking or alcohol consumption[107].

2.8 Other factors

This chapter is not intended to be a comprehensive review of the vast breast cancer epidemiology literature. There are a large number of other factors that have been considered in the literature such circadian disruption as a result of light at night exposure[108] and use of nonsteroidal anti-inflammatory drugs[109], among many others.

Chapter 3

UK BIOBANK AND METHODS

3.1 Introduction

The main objective of United Kingdom (UK) Biobank is to investigate the complex roles of lifestyle, environmental and genetic determinants of a wide range of health outcomes amongst men and women aged between 40 and 69 years at recruitment [110].

This chapter describes the methods of the UK Biobank and summarises the general approach to statistical analyses used in Chapters 4-9.

3.2 Design and methods

UK Biobank is a multi-centre, prospective, observational study with continuing follow-up of cause-specific morbidity and mortality. It aims to address some of the shortcomings of previous studies which include inadequate quantities of disease cases, incomplete or inadequate measures of potential risk factors and confound-

ing factors, and retrospective case-control study designs. A total of 503,317 participants were recruited, but 273,474 women were available for the data analyses (after men were excluded). The methods are described in detail elsewhere[110] and are summarised below.

Participants were recruited between 2006 and 2010 and attended a baseline assessment at one of the 22 assessment centres in England, Scotland, or Wales. UK Biobank baseline assessment included an extensive touchscreen questionnaire, covering sociodemographics, lifestyle and environment factors, family history, health and medical history, and sex-specific factors. Various physical measurements, including body size and composition, were also assessed during the visit. Biological samples were also collected at baseline and have been used for genotyping and measurement of biochemical markers on the whole cohort.

All participants are followed-up for various health outcomes via linkage to national electronic heath records. Repeat assessments of the baseline visit are conducted at regular intervals[110]. Further enhancements in UK Biobank include an accelerometer study on over 100,000 participants[111] and an imaging study currently accruing, with the aim to attain detailed magnetic resonance imaging scans on over 100,000 participants[112].

UK Biobank provides data in an anonymised form, which is encoded using an algorithm that is irreversible by researchers. UK Biobank maintains an internal database which allows only UK Biobank to reverse anonymise the data. In order to access the data, researchers must obtain approval of their applications from UK Biobank and are then subsequently able to download the requested encrypted datasets. As a part of the application process for data access, researchers are required to provide information on their data security systems and protocols and

UK Biobank has the right to audit the security systems[113].

3.3 Recruitment

Postal invitations were mailed to 9,238,453 individuals registered with the UK's National Health Service (NHS) who were between the ages of 40-69 years and lived within approximately 25 miles of a UK Biobank assessment centre. A total of 503,317 participants consented to join the UK Biobank study cohort and attended the baseline assessment centre visit between 2006 and 2010. The overall participation rate was 5.45%. The participation rate in women was 6.4%[114].

3.3.1 Data collection

At the UK Biobank assessment centre, participants visited the following stations in order: reception, touchscreen questionnaire, interview questionnaire and blood pressure, physical measurements, and sample collection.

The UK Biobank questionnaire was administered by computerised direct data entry. This touchscreen self-administered questionnaire took an average of 30 minutes to complete and was designed so that participants were asked questions deemed to be relevant based on personal characteristics and responses to previous questions. While completing the questionnaire, a check message would pop up if the participant entered extreme values or if no value was provided. One staff member was available to monitor and provide assistance to 10-12 participants at a time. Broad questionnaire topic areas include sociodemographics, lifestyle and environmental factors, early life factors, family history, psychosocial factors, health and medical history, sex-specific factors and cognitive function. The questions on

the UK Biobank touchscreen questionnaire that are relevant to this thesis are shown in Appendix A.

3.3.2 Baseline questionnaire

For information that was not easily captured by touchscreen questionnaire, such as responses that were not categorical or numerical or otherwise required detailed questioning, a subsequent computer-assisted personal interview (CAPI) was utilised. Participants were provided with a pre-aide memoire before attending the assessment centre to reduce the difficulty and inaccuracy of reporting information that may be difficult to recall during the visit such as medications, operations, family history, and birth details. Certain questions were asked in the interview only if the participant had given particular responses on the touchscreen questionnaire. For instance, if the participant indicated a particular medical condition on the touchscreen questionnaire, the interviewer would be prompted to ask the participant additional, specific questions about that specific condition. Structured search capabilities and pre-coded lists of diseases, drugs, and occupations were built into the CAPI system to improve efficiency and accuracy[110].

3.3.3 Body size and composition measures

At the physical measures station, participants were first asked to remove their shoes and socks before the measurement procedure began. Measurements relevant to analyses comprising this thesis are described here. The Seca-202 tape measure was used to measure standing and sitting height. The Wessex non-stretchable sprung tape measure was used to measure the circumference of the waist at the

umbilicus and circumference of the hip at the widest point. Body impedance was measured using the Tanita BC-418MA bio-impedance device. Participants who did not have a pacemaker or were not pregnant were asked to stand in bare feet on the measuring plate of the device while holding the handles firmly with their arms hanging loosely at their sides. An estimate of weight and estimated percent fat was directly recorded by the computer[110]. Dual-energy X-ray absorptiometry (DXA) was used to measure fat mass and percentage on a subset of participants beginning in 2014 using the GE-Lunar iDXA[115].

3.3.4 Blood sample collection and storage

The protocol for biological samples collection aimed to ensure standardised collection across assessment centres and to allow for the widest range of possible assays that could be pursued in future research. At the baseline assessment centre visit, 40-50 mL of blood was collected using a vacutainer system and blood was separated into tubes. The transport temperature for all tubes was 4°C, except for the acid citrate dextrose sample which was maintained at 18°C. Samples from each participant are stored in two geographically separate locations to protect against loss, a "working" archive with samples at -80°C that is typically first to be used for any research activities and a "back-up" archive with samples in liquid nitrogen vapour (-196°C)[110].

3.3.5 Genetic data

Approximately 50,000 participants were genotyped using the Affymetrix UK BiLEVE Axiom array. The other 450,000 participants were genotyped using the Affymetrix

UK Biobank Axiom® array, designed by the UK Biobank Array Design Group. The content of these two arrays have 95% overlap. The UK Biobank Axiom® array contains 820,967 SNP and indel markers, including 343 cancer common variants that were selected from a list of published common variants associated with cancer as identified by genome-wide association studies (GWAS). More specifically, the National Human Genome Research Institute GWAS catalogue and recent studies up to June 2013 were used. The content of the array is optimised for a population with European ancestry[116].

3.4 Ethics

UK Biobank has approval from the NHS North West Multicentre Research Ethics Committee, the National Information Governance Board for Health and Social Care in England and Wales, and the Community Health Index Advisory Group in Scotland. All participants provided written informed consent[110].

3.5 Follow-up

At enrolment, permission was obtained from participants to access all past and future medical records. To ensure that participants are not lost during follow-up, a number of identifiers were obtained such as the NHS number (in England and Wales) and the Community Health number (in Scotland). Other identifiers that were sought to allow linkage to other sources of health-related information and to help prevent loss to follow-up include name, date of birth, address, general practice, mobile telephone numbers, and email addresses[110]. Mortality data are provided

by NHS Digital (in England and Wales) and Information and Statistics Division, Scotland (in Scotland). Cancer registry data are provided by NHS Digital (in England and Wales) and National Records of Scotland, NHS Central Register (in Scotland)[117].

Participants contributed follow-up time from the date of recruitment until the earliest of any of the following outcomes: breast cancer diagnosis, diagnosis of another cancer (other than non-melanoma skin cancer), loss to follow-up, death, or end of follow-up for the study period.

3.5.1 Outcome of incident breast cancer

Breast cancer cases were defined by code C50, as specified by the 10th revision of the International Statistical Classification of Injuries, Disease and Deaths (ICD-10)[118]. Cases were defined as women who developed breast cancer after the date of recruitment and up until the end of follow-up.

3.5.2 Repeat assessments and additional measures at re-assessment

After baseline assessment at recruitment,repeat assessment was performed on 20,000 participants approximately 5 years after recruitment in the first instance and then again on those who have attended an imaging assessment. The imaging assessment is currently ongoing, for up to 100,000 participants.

Fluctuations in the measured baseline values of a risk factor tend to underestimate the strength of the real association between the long-term average level of the risk factor during a certain exposure period and incidence of the health outcome

of interest during the same or later period of time. Repeated measurements allow for correction of this regression dilution effect, which can be caused by biological variations in the short term (due to diurnal, seasonal, or otherwise transient variations) and/or longer-term individual variation in the risk factor (for example, due to factors such as diet, treatment, age, physical activity).

Collecting additional, more detailed information on a representative subset of the cohort allows for researchers to address imprecision in the baseline measures on the whole cohort, beyond adjusting for regression dilution. For example, a web-based 24-hour dietary recall questionnaire has been completed by a substantial portion of UK Biobank participants and provides more detailed information than the baseline food frequency questionnaire data available on the whole cohort[110].

3.6 Statistical analyses

Each individual results chapter includes a description of the specific methods used. This section outlines some of the common features of the statistical analyses.

For analysis of each given variable, women were excluded if they had missing data on that variable. Variables that were associated with both the exposure of interest and causally associated with the outcome were included in regression models as possible confounders. Other factors which have been considered as likely confounders by prior studies were also considered as covariates a priori, even if they did not appear to be associated with the exposure and were not a risk or preventive factor for the outcome in this dataset. For example, variables such as socioeconomic status, which may be a marker of a poorly understood set of factors related to breast cancer risk, were also treated as confounders in the analyses.

For cross-sectional analyses, multivariate linear regression models were used to examine the associations between risk factors for breast cancer. Cox proportional hazards models with attained age as the underlying time variable were used to estimate adjusted hazards ratios (referred to as relative risks (RRs)) of incident breast cancer according to the exposure of interest. Graphical methods and assessment of interaction terms between each exposure of interest and the underlying time variable were used to check for violation of the proportional hazards assumption.

All analyses were stratified by 5-year age at recruitment categories, geographical region of recruitment (London, Wales, North-West England, North-Eastern England, Yorkshire & the Humber, West Midlands, East Midlands, South-East England, South-West England, and Scotland), and socioeconomic status based on fifths of Townsend deprivation index[119] (assigned using postcode and based on preceding national census data). Adjustment variables are described in each of the individual results chapters. For each adjustment variable, missing values were assigned a separate category.

Chapter 4

RISK FACTORS FOR BREAST CANCER IN UK BIOBANK

4.1 Introduction

UK Biobank is an important new cohort for the study of risk factors for various health outcomes. This chapter describes the first results from this cohort on the association between breast cancer incidence and established risk factors. For completeness, analyses for BMI and physical activity are shown simply here but are investigated in greater detail in subsequent chapters.

4.2 Methods

The methods of recruitment and follow-up for UK Biobank have been described in detail in Chapter 3, and are summarised here.

4.2.1 Study design and participants

UK Biobank is a prospective cohort study of approximately half a million men and women between the ages of 40-69 at recruitment during 2006-2010. Information on data available and access procedures are described on the study website (http://www.ukbiobank.ac.uk/). At the time of the current analyses, the complete UK Biobank dataset included 502,620 participants (229,165 men and 273,455 women). During the baseline assessment centre visit, participants completed a touchscreen questionnaire which included questions on sociodemographics, lifestyle, health and medical history, and sex-specific factors. Physical measurements, such as body size and composition, were also assessed on the whole cohort during the visit. Women were excluded from the analyses described in this chapter if they had a prior cancer diagnosis (except for non-melanoma skin cancer ICD-10 C44) (n=18,372). A total of 255,083 women were included in the final analyses.

Women were defined as premenopausal or postmenopausal at recruitment based on whether they reported their periods had stopped; for these analyses, women with unknown self-reported menopausal status who were under the age of 45 who had self-reported having not undergone a bilateral oophorectomy were categorised as premenopausal and women age 53 or over and/or had had both ovaries removed were categorised as postmenopausal, based on previously established criteria[84]. A total of 60,745 premenopausal women and 179,621 postmenopausal women were included. After applying the above criteria, women with unknown menopausal status were excluded (n=14,717). A participant inclusion/exclusion flow diagram is shown in Figure 4.1.

UK Biobank was approved by the North West Multi-centre Research Ethics

Committee, the Confidentiality Advisory Group in England and Wales, and the Community Health Index Advisory Group in Scotland. All participants gave written informed consent.

4.2.2 Exposure assessment

Study participants provided self-reported responses using touchscreen questionnaires at the baseline assessment centre visit, along with a brief computer-assisted interview[120]. Questions from the touchscreen questionnaire relevant to the analyses presented in this thesis are shown in Appendix A. Standing height was measured using the Seca 202 device (Seca, Hamburg, Germany). BMI was calculated by dividing weight (kg) by the square of standing height (m^2).

Exposures investigated in this chapter include sociodemographic factors, anthropometric factors, lifestyle factors, reproductive factors, breast cancer screening (ever had mammogram) and family history.

The sociodemographic factors include region (London, Wales, North-West England, North-Eastern England, Yorkshire & the Humber, West Midlands, East Midlands, South-East England, South-West England, and Scotland), socioeconomic status (based on fifths of Townsend deprivation index), ethnicity (White, Mixed, Asian, Black, Chinese, Other), educational qualifications (college or university, A/AS levels or equivalent, O levels/GCSE or equivalent, CSEs or equivalent, NVQ or HND or HNC or equivalent, other professional qualifications, none of the above), and employment status (in paid employment or self-employed, retired, looking after home and/or family, unable to work because of sickness or disability, unemployed, doing paid or voluntary work, full or part-time student, none of the

above). For the current analyses, O levels/GSCEs and CSEs were grouped together. NVQ or HND or HNC were grouped with other professional qualifications to create a vocational qualifications category. For employment status, categories other than "in paid employment or self-employed" and "retired" were grouped into an "other" category due to small numbers in these groups.

Anthropometric factors included height in cm (grouped into the following categories: <160, 160-163.9, 164.0-167.9, ≥168.0), BMI in kg/m^2 (grouped as <25, 25-29, ≥30), and comparative body size at age 10 (thinner, about average, plumper).

Lifestyle factors included physical activity (grouped as low, moderate, high), smoking status (never, past, current), and alcohol intake in grams per day (grouped as non-drinkers, <1, 1-10, ≥10). Physical activity was comprised of the sum of walking (3.3 metabolic equivalents [METs]), moderate activity (4.0 METs), and vigorous activity (8.0 METs). Low physical activity ranged from 0-10 MET-hours per week, moderate ranged from 10-49.9 MET-hours per week, and high was ≥50 MET-hours per week. Further details on baseline physical activity variables are described in depth in Chapter 5.

Amount of alcohol consumption in grams was calculated by summing the units consumed of the following types of drinks: red wine, champagne plus white wine, beer plus cider, spirits, fortified wine, and other. One glass of any type of wine or unit of spirits was considered to be 10 grams of alcohol. A pint of beer or cider was considered to be 20 grams of alcohol. Non-drinkers were identified by responses to the question, "About how often do you drink alcohol?"

Reproductive factors included age at menarche in years (grouped as <12, 12-13, ≥14), oral contraceptive use (never, past, current), age at menopause in years (grouped as <48, 48-52, 53-55, ≥56), and hormone replacement therapy (HRT)

use (never, past, current). Since parity and age at first birth are related, a cross-classified variable was created with the following categories: nulliparous, 1-2 births and <25 years, 1-2 births and ≥25 years, ≥3 births and <25 years, ≥3 births and ≥25 years.

For breast cancer screening, participants were asked if they had ever been for breast cancer screening (a mammogram). Family history of breast cancer was defined as breast cancer diagnosis in mother or a sister. Family history of breast cancer was not available for daughters.

4.2.3 Ascertainment of cancer cases

UK Biobank obtains data on cancer diagnoses through NHS Digital for participants in England and Wales, and NHS Central Register for participants in Scotland.

4.2.4 Statistical analysis

Women were followed from date of baseline assessment centre visit until the earliest of: date of breast cancer registration (ICD-10 C50), date of death, date of loss to follow-up or end of follow-up for cancer incidence (30 December 2014). Women diagnosed with any cancer other than breast cancer (with the exception of non-melanoma skin cancer) during follow-up were censored at date of diagnosis.

Cox proportional hazards models with attained age as the underlying time variable were used to calculate hazard ratios (referred to as relative risks (RRs)) and 95% confidence intervals (CIs) for breast cancer risk. Analyses were stratified by 5-year age at recruitment categories, region of recruitment, and socioeconomic status (based on fifths of Townsend deprivation index[119]) except for when the variable

was the main exposure of interest. Exposures and covariates were selected a priori based on known risk factors for breast cancer. Multivariable-adjusted analyses were adjusted for height, comparative body size at age 10, physical activity level, smoking status, alcohol intake, family history of breast cancer, ever had mammogram, age at menarche, age at first birth, parity, oral contraceptive use, hormone replacement therapy use, age at menopause, and BMI when the covariate was not the main exposure of interest. As a covariate, the association of BMI with risk was allowed to vary by menopausal status by adjusting for a variable generated by the cross-classification of BMI and menopausal status. The analyses for BMI were performed separately in premenopausal and postmenopausal women. Missing and unknown values for the exposure of interest were not included in the analyses but were assigned to a separate unknown category when the variable was included as a covariate.

P-values from the multivariable-adjusted models for the association of each exposure of interest with breast cancer risk were calculated in the following ways: for variables with more than two categories, P-values were calculated using a likelihood ratio test comparing the model with and without the exposure of interest; for dichotomous variables, P-values were calculated using the reference category as the comparison; for continuous variables, P-values for trend were calculated using a pseudo-continuous variable recoded with the median value in each category. Assessment of interaction terms between each exposure of interest and the underlying time variable did not suggest any significant deviation from proportional hazards. All analyses were conducted using STATA version 15.0 (Stata Corp, LP, College Station, TX).

4.3 Results

Among a total of 255,083 women included in the final analyses, 4,522 women were diagnosed with breast cancer during a mean 5.8 years (SD 1.1 years) follow-up. Table 4.1 shows the baseline characteristics of the study population. The mean age at recruitment was 56.1 years (SD 8.0 years). The mean height was 162.4 cm (SD 6.3 cm), the mean BMI was 27.1 kg/m^2 (SD 5.2 kg/m^2), and 17.5% of women reported being plumper than average at age 10. On average, participants did 29.4 MET-hours a week of physical activity, which is in the moderate activity range. Nine percent of participants reported being current smokers and 35.4% reported drinking at least 10 grams of alcohol per day. Overall, 6.3% of participants had a family history of breast cancer and 78.4% had ever had a mammogram. Mean ages at menarche and menopause were 13.0 years (SD 1.6 years) and 49.7 years (SD 5.1 years), respectively. Mean age at first birth was 25.9 years (SD 5.1 years) and 18.6% of women were nulliparous. Nearly 2% were current users of oral contraceptives and 7.6% were current users of hormone replacement therapy.

Relative risks of breast cancer in relation to various sociodemographic factors are shown in Table 4.2, anthropometric and lifestyle factors are shown in Table 4.3, and reproductive and other factors are shown in Table 4.4. Relative risks from models that only stratified by age, region, and socioeconomic status were similar to relative risks from multivariable-adjusted models.

4.3.1 Anthropometric factors

Height per 10 cm increase was associated with an increased risk of breast cancer (RR 1.20, 95% CI 1.13-1.27). BMI was associated with breast cancer risk in

postmenopausal women (RR 1.41; 95% CI 1.29-1.55) but not in premenopausal women (RR 0.94, 95% CI 0.79-1.12), per 10 kg/m^2 increase. Being plumper than average at age 10 was associated with decreased risk (RR 0.80, 95% CI 0.73-0.88) while being thinner than average was associated with increased risk (RR 1.08, 95% CI 1.01-1.16) (Table 4.3). Further analyses of BMI in postmenopausal women are shown in Chapter 8.

4.3.2 Lifestyle factors

Alcohol intake of 10 grams or more per day was associated with a 14% increased risk of breast cancer (RR 1.14; 95% CI 1.04-1.26), compared to alcohol intake of $<$1 gram per day. Physical activity was protective (RR 0.84, 95% CI 0.76-0.94), comparing the most active to least active women. Smoking status overall was not associated with breast cancer risk, but the relative risk in current smokers was borderline significant (RR 1.11, 95% CI 1.00-1.25) (Table 4.3). Further analyses of physical activity in relation to breast cancer risk in premenopausal and postmenopausal women are shown in Chapter 9.

4.3.3 Reproductive factors

Older age at menarche was protective ($\geq$14 years vs 12-13 years: RR 0.91, 95% CI 0.85-0.97). Women with 3 or more births and age at first birth younger than 25 years had a 19% reduced risk (RR 0.81; 95% CI 0.72-0.90), compared to nulliparous women. A nearly identical risk reduction was seen in women with 1-2 births and age at first birth younger than 25 years. Oral contraceptive use (current vs never: RR 1.69, 95% CI 1.38-2.08), hormone replacement therapy use (current vs never:

RR 1.68, 95% CI 1.49-1.89), and older age at menopause (≥56 years vs 48-52 years: RR 1.19, 95% CI 1.04-1.36) were all associated with increased breast cancer risk (Table 4.4).

4.3.4 Other factors

Having never had a mammogram was associated with a small, borderline reduced risk for breast cancer (RR 0.89, 95% CI 0.79-1.00). Having a family history of breast cancer in the mother and/or sister was associated with increased risk (RR 1.61, 95% CI 1.46-1.78) compared to no family history of breast cancer (Table 4.4).

4.4 Discussion

These analyses in a large prospective cohort study of 4,522 incident cases of breast cancer in women confirm many of the established breast cancer risk factors[2,121]. Height, BMI for postmenopausal women, and being comparatively thinner at age 10 were associated with increased breast cancer risk. Breast cancer risk was lower in those who were more physically active and higher in those who consumed more alcohol. Older age at menarche as well as a combination of greater number of births and younger age at first birth were protective. Current use of oral contraceptives, older age at menopause, current use of hormone replacement therapy, and family history of breast cancer were associated with increased risk.

Findings from the present study of a 20% increased risk of breast cancer per 10 cm increase in measured height are consistent with results from a 2015 meta-analysis that found a 17% increased risk in all studies combined and a 20% increased risk among studies with measured height, per 10 cm increase[18]. A 10 unit

increase in BMI in postmenopausal women was associated with a 41% increase in risk, which is in agreement with results from the Million Women Study[21] of a 40% increase in breast cancer risk per 10 kg/m^2. The estimates presented in this chapter may be expected to be of smaller magnitude given that HRT users were not excluded. However, the fact that BMI was measured rather than self-reported in UK Biobank may result in a stronger association. There was no significant association observed for BMI and breast cancer risk in premenopausal women. A recent umbrella review concluded that the evidence for an inverse association between adiposity and premenopausal breast cancer remains suggestive[122]. However, it appears to me that the evidence for an inverse association between adiposity and premenopausal breast cancer is stronger, and many of the analyses, including the data presented here, are likely to have been underpowered to detect an association. For childhood adiposity, the results presented here demonstrating a protective association for those who reported being comparatively plumper at age 10 are consistent with prior studies[34,123].

The 16% reduction in risk for the high compared to low physical activity group observed in the present analyses is in agreement with prospective cohort estimates from a recent large meta-analysis that found a 17% reduced breast cancer risk for premenopausal women and a 10% reduced risk for postmenopausal women[44].

Whether there exists a causal relationship between cigarette smoking and breast cancer risk remains controversial and was deemed "suggestive but not sufficient" by the 2014 US Surgeon General's report[124] and inconclusive by the International Agency for Research on Cancer due to limited evidence to support or refute carcinogenicity[63]. There was a borderline significant 11% increased risk in current compared to never smokers, which is consistent with the 7% increased risk

in current smokers observed in a pooled analysis in the National Cancer Institute Cohort Consortium[65]. Smoking status overall was not associated with breast cancer risk in the present study. However, an association cannot be ruled out in certain subgroups since the analyses shown here were not stratified for some of the characteristics that have been shown to be most strongly associated with risk: smoking initiation more than 10 years before first birth, consumption of two or more alcoholic drinks per day, and ER^+ breast cancer receptor status[65].

For alcohol intake, the results from a recent pooled analysis reporting a 10% increased breast cancer risk in women who consumed 5-15 grams of alcohol per day compared to non-drinkers[58] are in line with the results shown in this chapter of a 6% increased breast cancer risk per 10 grams of alcohol consumed per day.

For reproductive factors, findings in UK Biobank are consistent with prior findings. In line with a large pooled analysis reporting a 5% increased risk for every year younger at menarche and a 3% increased risk for every year older at menopause[67], the analyses in this chapter demonstrate a 3% reduced risk per every year older at menarche and a 2% increased risk for every year older at menopause. Compared to nulliparous women, the results here show that younger age at first birth and greater parity were associated with lower risk and are consistent with the prior literature[73,125].

Findings for the associations between use of exogenous hormones and breast cancer risk are also consistent with the literature. The results for oral contraceptive use are broadly similar to those reported from a large pooled analysis, which found a 24% increased risk in current users and no excess risk 10 or more years after stopping use[79]. In the analysis presented here, the greater magnitude of increased risk observed in women who had reported current use of oral contraceptives is

likely explained by chance since a relatively small number of women developed breast cancer within that group. The finding reported here of a 68% increased risk of breast cancer in current users compared to never users of HRT at baseline is consistent with the 66% increased risk reported in the Million Women Study[84].

The results shown here of a 61% increased risk for women with a family history of breast cancer is consistent with prior estimates[101]. There was a slightly decreased risk in women who had never had a mammogram, which may reflect some degree of overdiagnosis[126].

To my knowledge, this is the first comprehensive study of risk factors for breast cancer in UK Biobank. Strengths include the prospective design, large sample size, virtually complete follow-up[127], and the availability of data on a wide range of potential confounders. However, the possibility of residual confounding cannot be excluded. Information on some exposures, such as breastfeeding, mammographic density, and benign breast disease, was not available. Moreover, detailed information on tumour characteristics, such as stage, grade, and hormone receptor status, was not available at the time of writing of this thesis. Although the number of missing values for the exposures investigated in this chapter is generally low (<2%), there are a few variables in this analysis with a higher proportion of missing values at baseline assessment: physical activity (26.2%), age at menarche (3.3%), and age at menopause (20.7%). These values may not be missing at random. For example, age at menopause can be difficult to report accurately since menopausal status may be masked by hysterectomy or HRT use before menstrual periods naturally stop[84]. Variables with a large percentage of values that are missing not at random may bias the results of the analyses. This bias may be a problem in both analyses based on analyses of complete cases and in analyses based on multiple imputation

(not used in this thesis). Issues around self-reporting of physical activity will be discussed in later chapters.

4.5 Conclusion

In this chapter, I have reported on a wide range of established risk factors for breast cancer in UK Biobank and demonstrated that the results are consistent with those reported in other prospective studies. Future studies in UK Biobank will be able to assess the associations between the exposures investigated here and breast cancer risk by tumour subtype.

Table 4.1: Baseline characteristics of all women included in the analysis and of women who developed breast cancer

Baseline characteristics	**All women**	**Women who developed breast cancer**
Number of women	255,083	4,522
Age at recruitment (years), mean (SD)	56.1 (8.0)	57.3 (7.7)
Lowest fifth of socioeconomic status (%)	51,194 (20.1%)	774 (17.1%)
White (%)	239,761 (94.0%)	4,338 (95.9%)
College or university degree (%)	79,334 (31.1%)	1,431 (31.6%)
Employment status		
In paid employment or self-employed (%)	141,827 (55.6%)	2,378 (52.6%)
Retired (%)	86,373 (33.9%)	1,728 (38.2%)
Other (%)	26,462 (10.4%)	409 (9.0%)
Height, mean (SD)	162.4 (6.3)	162.9 (6.5)
BMI, mean (SD)	27.1 (5.2)	27.4 (5.1)
Plumper than average at age 10 (%)	44,718 (17.5%)	659 (14.6%)
Physical activity (MET-hours/week), median (IQR)	29.4 (13.6, 58.2)	27.1 (12.6, 53.4)
Current smoker (%)	22,849 (9.0%)	398 (8.8%)
Alcohol intake ≥10 g per day, (%)	90,290 (35.4%)	1,709 (37.8%)
Age at menarche (years), mean (SD)	13.0 (1.6)	12.9 (1.6)
Age at first birth (years), mean (SD)	25.9 (5.1)	26.2 (5.3)
Nulliparous (%)	47,554 (18.6%)	872 (19.3%)
Current use of oral contraceptives (%)	4,858 (1.9%)	116 (2.6%)
Age at menopause (years), mean (SD)	49.7 (5.1)	50.3 (5.0)
Current use of hormone replacement therapy (%)	13680 (7.6%)	370 (10.9%)
Ever had mammogram (%)	199,993 (78.4%)	3,752 (83.0%)
Family history of breast cancer (%)	16,174 (6.3%)	456 (10.1%)

Abbreviations: body mass index (BMI), interquartile range (IQR), metabolic equivalent (MET), standard deviation (SD)
Percentages may not sum exactly to 100% due to missing data.

Notes for Tables 4.2, 4.3, and 4.4:

Abbreviations: body mass index (BMI), confidence interval (CI), metabolic equivalent (MET), relative risk (RR).

[a] Stratified by age at recruitment, region of recruitment, and socioeconomic status (Townsend deprivation index)

[b] Adjusted for height, comparative body size at age 10, physical activity, smoking status, alcohol intake, family history of breast cancer, ever had mammogram, age at menarche, age at first birth, parity, oral contraceptive use, hormone replacement therapy use (in postmenopausal women), age at menopause, and BMI (within each menopausal status group)

[c] For variables with more than two categories, p-values were calculated using a likelihood ratio test comparing the model with and without the variable of interest. For dichotomous variable, p-values were calculated using the reference category as the comparison. For continuous variables, p-values for trend were calculated using a pseudo-continuous variable recoded with the median value in each category

Table 4.2: Relative risks (95% CI) for breast cancer in relation to sociodemographic factors at baseline

Characteristics	Total n	Cases	RR (95% CI) stratified only [a]	RR (95% CI) multivariable-adjusted [b]	P-value [c]
Sociodemographic factors					
Region					0.13
London	35,898	509	1 (reference)	1 (reference)	
Wales	10,504	194	0.97 (0.82-1.14)	1.01 (0.85-1.19)	
North-West England	38,969	770	1.08 (0.97-1.22)	1.13 (1.01-1.27)	
North-Eastern England	29,629	521	1.01 (0.89-1.14)	1.06 (0.94-1.20)	
Yorkshire & the Humber	37,740	591	0.94 (0.83-1.06)	0.98 (0.87-1.10)	
West Midlands	21,495	352	1.02 (0.89-1.17)	1.07 (0.93-1.23)	
East Midlands	17,283	291	0.96 (0.83-1.12)	1.01 (0.87-1.17)	
South-East England	22,521	472	1.08 (0.95-1.22)	1.09 (0.95-1.24)	
South-West England	22,385	421	1.11 (0.97-1.27)	1.15 (1.00-1.31)	
Scotland	18,659	401	1.08 (0.94-1.23)	1.12 (0.98-1.28)	
Socioeconomic status (based on Townsend deprivation index), fifths					0.033
One (most affluent)	50,852	982	1 (reference)	1 (reference)	
Two	50,829	933	0.96 (0.88-1.05)	0.97 (0.88-1.06)	
Three	50,937	904	0.94 (0.86-1.03)	0.94 (0.86-1.03)	
Four	50,964	926	0.99 (0.90-1.08)	1.00 (0.91-1.10)	
Five (most deprived)	51,194	774	0.84 (0.76-0.92)	0.87 (0.78-0.96)	
Ethnicity					0.093
White	239,761	4,338	1 (reference)	1 (reference)	
Mixed	1,757	28	1.04 (0.72-1.52)	1.07 (0.74-1.56)	
Asian	4,428	56	0.84 (0.64-1.09)	1.03 (0.78-1.36)	
Black	4,496	36	0.55 (0.39-0.77)	0.63 (0.45-0.88)	
Chinese	952	10	0.65 (0.35-1.21)	0.78 (0.42-1.46)	
Other	2,477	33	0.88 (0.62-1.24)	1.00 (0.70-1.41)	
Education					0.457
College or university	79,334	1,431	1 (reference)	1 (reference)	
A levels or equivalent	29,995	529	0.97 (0.88-1.07)	0.99 (0.90-1.10)	
GCSE or equivalent	72,713	1,235	0.92 (0.86-1.00)	0.97 (0.90-1.05)	
Vocational qualifications	25,868	425	0.86 (0.77-0.96)	0.91 (0.82-1.02)	
None of the above	42,214	799	0.93 (0.85-1.02)	1.01 (0.92-1.12)	
Employment status					0.448
Paid/self-employed	141,827	2,378	1 (reference)	1 (reference)	
Retired	86,373	1,728	0.93 (0.85-1.01)	0.94 (0.86-1.03)	
Other	26,462	409	0.95 (0.86-1.06)	1.00 (0.89-1.11)	

Notes for Tables 4.2, 4.3, and 4.4:

Abbreviations: body mass index (BMI), confidence interval (CI), metabolic equivalent (MET), relative risk (RR).

[a] Stratified by age at recruitment, region of recruitment, and socioeconomic status (Townsend deprivation index)

[b] Adjusted for height, comparative body size at age 10, physical activity, smoking status, alcohol intake, family history of breast cancer, ever had mammogram, age at menarche, age at first birth, parity, oral contraceptive use, hormone replacement therapy use (in postmenopausal women), age at menopause, and BMI (within each menopausal status group)

[c] For variables with more than two categories, p-values were calculated using a likelihood ratio test comparing the model with and without the variable of interest. For dichotomous variable, p-values were calculated using the reference category as the comparison. For continuous variables, p-values for trend were calculated using a pseudo-continuous variable recoded with the median value in each category

Table 4.3: Relative risks (95% CI) for breast cancer in relation to anthropometric and lifestyle factors at baseline

Characteristics	Total n	Cases	RR (95% CI) stratified only [a]	RR (95% CI) multivariable-adjusted [b]	P-value [c]
Anthropometric factors					
Height, cm					
<160	81,365	1,359	1 (reference)	1 (reference)	
160-163.9	63,597	1,054	1.02 (0.94-1.10)	1.02 (0.94-1.10)	
164.0-167.9	56,377	1,034	1.15 (1.06-1.25)	1.14 (1.05-1.24)	
≥168.0	52,666	1,057	1.31 (1.21-1.42)	1.29 (1.19-1.40)	
Per 10 cm increase	254,005	4,504	1.21 (1.14-1.29)	1.20 (1.13-1.27)	<0.0001
BMI (premenopausal)					
<25	29,091	466	1 (reference)	1 (reference)	
25-29	18,973	284	0.93 (0.80-1.08)	0.96 (0.83-1.12)	
≥30	12,430	168	0.87 (0.73-1.04)	0.94 (0.78-1.13)	
Per 10 kg/m^2 increase	60,494	918	0.88 (0.74-1.03)	0.94 (0.79-1.12)	0.466
BMI (postmenopausal)					
<25	67,292	1,118	1 (reference)	1 (reference)	
25-29	67,406	1,328	1.18 (1.09-1.27)	1.23 (1.13-1.33)	
≥30	44,122	923	1.28 (1.17-1.40)	1.42 (1.29-1.56)	
Per 10 kg/m^2 increase	178,820	3,369	1.27 (1.17-1.39)	1.41 (1.29-1.55)	<0.0001
Body size at age 10					<0.0001
Thinner	79,500	1,510	1.06 (1.00-1.14)	1.08 (1.01-1.16)	
About average	126,099	2,271	1 (reference)	1 (reference)	
Plumper	44,718	659	0.84 (0.77-0.91)	0.80 (0.73-0.88)	
Lifestyle factors					
Physical activity, MET-hours/week					
Low	33,913	651	1 (reference)	1 (reference)	
Moderate	97,225	1,733	0.92 (0.84-1.00)	0.92 (0.84-1.01)	
High	56,055	891	0.81 (0.74-0.90)	0.84 (0.76-0.94)	
Per 20 METs/week increase	187,193	3,275	0.87 (0.81-0.93)	0.89 (0.83-0.96)	0.002
Smoking status					0.126
Never	151,927	2,620	1 (reference)	1 (reference)	
Past	78,906	1,481	1.07 (1.00-1.14)	1.04 (0.97-1.11)	
Current	22,849	398	1.08 (0.97-1.20)	1.11 (1.00-1.25)	
Alcohol intake, grams/day					
Non-drinkers	24,218	393	1.02 (0.90-1.16)	1.05 (0.92-1.19)	
<1 g/d	38,594	618	1 (reference)	1 (reference)	
1-10 g/d	99,891	1,774	1.10 (1.01-1.21)	1.09 (1.00-1.20)	
≥10 g/d	90,290	1,709	1.17 (1.07-1.29)	1.14 (1.04-1.26)	
Per 10 g per day increase	252,993	4,494	1.08 (1.03-1.12)	1.06 (1.01-1.10)	0.013

Notes for Tables 4.2, 4.3, and 4.4:

Abbreviations: body mass index (BMI), confidence interval (CI), metabolic equivalent (MET), relative risk (RR).

[a] Stratified by age at recruitment, region of recruitment, and socioeconomic status (Townsend deprivation index)

[b] Adjusted for height, comparative body size at age 10, physical activity, smoking status, alcohol intake, family history of breast cancer, ever had mammogram, age at menarche, age at first birth, parity, oral contraceptive use, hormone replacement therapy use (in postmenopausal women), age at menopause, and BMI (within each menopausal status group)

[c] For variables with more than two categories, p-values were calculated using a likelihood ratio test comparing the model with and without the variable of interest. For dichotomous variable, p-values were calculated using the reference category as the comparison. For continuous variables, p-values for trend were calculated using a pseudo-continuous variable recoded with the median value in each category

Table 4.4: Relative risks (95% CI) for breast cancer in relation to reproductive factors and other factors at baseline

Characteristics	Total n	Cases	RR (95% CI) stratified only [a]	RR (95% CI) multivariable-adjusted [b]	P-value [c]
Reproductive factors					
Age at menarche, years					
<12	49,252	911	1.00 (0.92-1.08)	1.02 (0.94-1.10)	
12-13	106,894	1,960	1 (reference)	1 (reference)	
≥14	90,626	1,530	0.92 (0.86-0.98)	0.91 (0.85-0.97)	
Per 1 year increase	246,772	4,401	0.97 (0.96-0.99)	0.97 (0.95-0.98)	<0.0001
Parity and age at first birth					<0.0001
Nulliparous	47,554	872	1 (reference)	1 (reference)	
1-2 births and <25 years	50,389	839	0.81 (0.73-0.89)	0.81 (0.73-0.89)	
1-2 births and ≥25 years	94,508	1,761	0.95 (0.87-1.03)	0.95 (0.87-1.03)	
3+ births and <25 years	36,167	607	0.81 (0.73-0.90)	0.81 (0.72-0.90)	
3+ births and ≥25 years	25,244	423	0.84 (0.74-0.94)	0.84 (0.74-0.94)	
Oral contraceptive use					<0.0001
Never	47,354	862	1 (reference)	1 (reference)	
Past	201,516	3,526	1.04 (0.96-1.13)	1.01 (0.93-1.09)	
Current	4,858	116	1.78 (1.46-2.19)	1.69 (1.38-2.08)	
Age at menopause, years					
<48	35,142	572	0.92 (0.83-1.02)	0.89 (0.80-0.99)	
48-52	67,200	1,218	1 (reference)	1 (reference)	
53-55	28,151	589	1.12 (1.01-1.24)	1.12 (1.01-1.24)	
≥56	12,077	281	1.19 (1.05-1.36)	1.19 (1.04-1.36)	
Per 1 year increase	142,520	2660	1.02 (1.01-1.03)	1.02 (1.01-1.03)	<0.0001
Hormone replacement therapy use (postmenopausal)					<0.0001
Never	87,747	1,505	1 (reference)	1 (reference)	
Past	77,587	1,495	1.03 (0.95-1.11)	1.05 (0.97-1.13)	
Current	13,680	370	1.63 (1.45-1.83)	1.68 (1.49-1.89)	
Other factors					
Ever had mammogram					0.045
No	54,133	755	0.86 (0.77-0.97)	0.89 (0.79-1.00)	
Yes	199,993	3,752	1 (reference)	1 (reference)	
Family history of breast cancer (in mother or sister)					<0.001
No	234,843	4,066	1 (reference)	1 (reference)	
Yes	15,718	456	1.63 (1.48-1.80)	1.61 (1.46-1.78)	

Notes for Tables 4.2, 4.3, and 4.4:

Abbreviations: body mass index (BMI), confidence interval (CI), metabolic equivalent (MET), relative risk (RR).

[a] Stratified by age at recruitment, region of recruitment, and socioeconomic status (Townsend deprivation index)

[b] Adjusted for height, comparative body size at age 10, physical activity, smoking status, alcohol intake, family history of breast cancer, ever had mammogram, age at menarche, age at first birth, parity, oral contraceptive use, hormone replacement therapy use (in postmenopausal women), age at menopause, and BMI (within each menopausal status group)

[c] For variables with more than two categories, p-values were calculated using a likelihood ratio test comparing the model with and without the variable of interest. For dichotomous variable, p-values were calculated using the reference category as the comparison. For continuous variables, p-values for trend were calculated using a pseudo-continuous variable recoded with the median value in each category

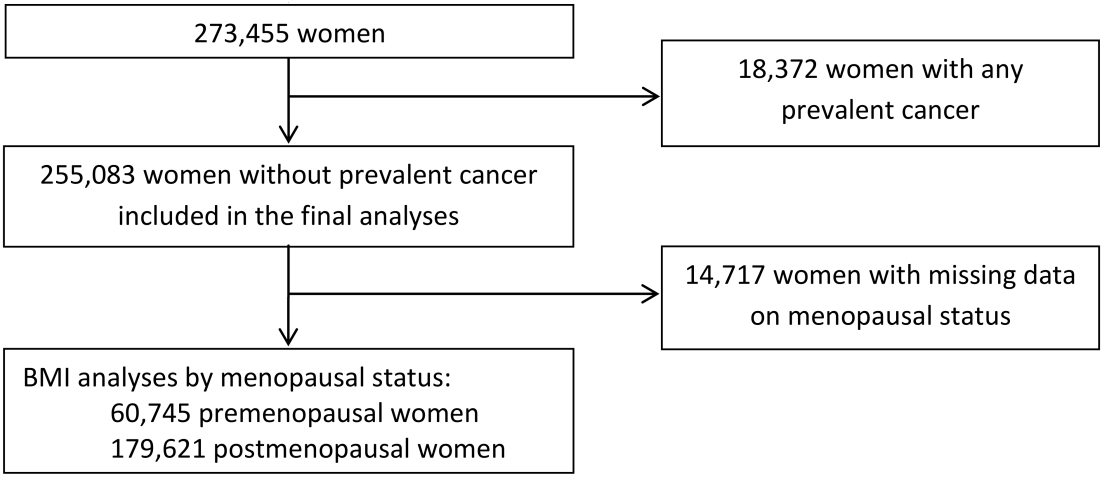


Figure 4.1: Participants inclusion/exclusion flow diagram for the analyses of risk factors for breast cancer in UK Biobank

Chapter 5

SELF-REPORTED PHYSICAL ACTIVITY IN RELATION TO BODY SIZE AND COMPOSITION

5.1 Introduction

Obesity is of major, general public health concern due to its association with increased risk of a variety of diseases including diabetes, heart disease, hypertension, stroke, and certain cancers, such as breast and endometrial[128]. Although physical activity is widely believed to be an important determinant of body size and composition, the full nature of the relationship is still unclear[129,130]. While observational studies support an association between exercise and weight-loss maintenance, randomised controlled trials have produced inconsistent results[8]. Questions also remain over the role of vigorous compared with lower intensity physical activity, and over the relationship between physical activity and different measures of overall

versus central adiposity[131,132]. Since the association between adiposity and breast cancer risk differs by menopausal status, it is of interest to examine the associations between physical activity, adiposity, and menopausal status in women[2,133]. Some previous large studies have been limited by low numbers of women reporting vigorous physical activity[134].

A number of anthropometric variables have been reported and utilised in public health recommendations. Since an increase in particular indices of body size may have an impact on specific health outcomes, it is important to assess which indices are likely to be modified by exercise. This chapter reports on a cross-sectional analysis of associations between reported levels of overall and of vigorous physical activity and several objective measures of body size and composition in over 132,000 pre- and postmenopausal women. In particular, I assess the hypothesis that physical activity is related to body size and composition and that associations may vary by particular measures of adiposity.

5.2 Methods

5.2.1 Data source

Data were obtained by application to UK Biobank (reference number 3248, approved August 2013). Details of the rationale, design and survey methods for UK Biobank have been described elsewhere[110] and information on data available and access procedures are given on the study website (http://www.ukbiobank.ac.uk/). UK Biobank has approval from the North West Multi-centre Research Ethics Committee, the National Information Governance Board for Health & Social Care in

England and Wales, and the Community Health Index Advisory Group in Scotland. All participants provided written informed consent.

5.2.2 Study participants

The complete UK Biobank dataset at the time these analyses were run included up to 502,656 participants (229,182 men and 273,474 women) between the ages of 40-69 at recruitment during 2006-2010.

A participant inclusion/exclusion flow diagram for this chapter is shown in Figure 5.1. Women were excluded from the analyses if they reported a long-standing illness/disability/infirmity or had unknown status for this variable (n=87,989), were pregnant (n=121), or had missing data on any of the physical activity variables used in the analyses (n=43,390). Participants were also excluded if the sum of all walking, moderate and vigorous time variables was greater than 6,720 minutes per week as adapted from the International Physical Activity Questionnaire (IPAQ) scoring guidelines (n=1,487), which are shown in Appendix B. After applying the above criteria for unknown menopausal status, menopausal status remained unknown for 7,694 women, who were excluded. Up to 38,201 premenopausal and 94,592 postmenopausal women were included in the final analyses.

5.2.3 Physical activity

Physical activity questions asked on the baseline questionnaire captured the frequency and duration of three different levels of activity (walking, moderate, and vigorous). Participants were asked how many days per typical week they spent engaged in each category of activity. For each category in which an answer of one or

more days was given, the participant was then asked the number of minutes spent on the activity during a typical day. Questions were adapted from the IPAQ short form, a validated survey instrument[135]. Questions are shown in Appendix A and are available to view on the UK Biobank website (http://biobank.ctsu.ox.ac.uk/). Metabolic equivalents (METs) were used to quantify physical activity; 1 MET is expended by sitting quietly for 1 hour, and the MET value reflects the ratio of energy expended per kilogram of body weight per hour to that expended when sitting quietly[136]. The number of minutes per day engaged in each level of activity was multiplied by the respective MET score for that level of activity. MET minutes per day were converted to MET hours per week. The total amount of METs was calculated by summing total METs from each activity level. In line with IPAQ guidelines[135], physical activity of less than 10 minutes per day for walking, moderate, or vigorous activity was recoded to 0. This was done in order to stay consistent with other studies that have used the IPAQ guidelines, which are widely used.

Each category was assigned the following MET hour values: 3.3 for walking, 4.0 for moderate physical activity, and 8.0 for vigorous physical activity[135]. Since individuals are estimated to expend 1 MET per hour sitting quietly, I report physical activity as excess METs[137]. Excess METs were calculated by subtracting 1 from the MET value multiplier for each category, giving values of 2.3 for walking, 3.0 for moderate physical activity, and 7.0 for vigorous physical activity. Please note that excess METs are used only in the present chapter to remain consistent with the paper published from the work in this chapter. In subsequent chapters, normal METs are used.

To assess the importance of vigorous physical activity relative to total physical

activity, vigorous physical activity was divided by total physical activity to generate a ratio of vigorous-to-total physical activity. The median ratio of vigorous-to-total physical activity in participants who engaged in any vigorous physical activity was 33%. For analysis, participants were grouped into three categories of vigorous-to-total physical activity: 0 vigorous physical activity, <33% vigorous physical activity, and >33% vigorous physical activity.

5.2.4 Anthropometry and body composition

At the UK Biobank baseline interview, trained staff used standard procedures to measure and record standing height using the Seca 202 device. The Wessex non-stretchable sprung tape measure was used to measure waist and hip circumference. The Tanita BC-418MA body composition analyser was used to assess bioimpedance measures including body fat percentage, whole body fat mass, trunk fat percentage, and trunk fat mass. UK Biobank derived BMI by dividing weight (kg) by the square of standing height (m^2). Waist-to-hip ratio (WHR) was derived by dividing waist circumference by hip circumference.

5.2.5 Statistical analyses

Multivariable linear regression models were used to examine the association between physical activity and the various anthropometric and body composition measures. Amount of physical activity was categorised into nine groups to obtain an approximately even distribution of participants (0, 1-4.9, 5-9.9, 10-14.9, 15-24.9, 25-34.9, 35-49.9, 50-74.9, and ≥75 excess MET h/week) and the figures show the mean within each category. Those who did no physical activity were placed in a

separate category even though it contained fewer individuals. Participant characteristics are reported for three broad levels of physical activity: low (<10 excess MET-h/week), moderate (10-49.9 excess MET-h/week), and high (≥50 excess MET-h/week). The relationship between total physical activity and anthropometric variables was examined separately for those for whom all physical activity was non-vigorous, less than 33% of total physical activity was vigorous, and more than 33% of total physical activity was vigorous. Ratios of vigorous-to-total physical activity excess MET-hours per week were derived and separate regression models were used for each category of vigorous-to-total physical activity. All analyses were performed separately for premenopausal and postmenopausal women. Analyses were conducted using STATA version 13.0 (State Corp LP, College Station, TX).

Covariates included 5- year age at recruitment categories, educational qualifications, parity, smoking status, employment status, alcohol intake, and fifths of Townsend deprivation index[119]. Educational qualifications were classified as vocational qualifications (other professional qualifications/NVQ or HND or HNC or equivalent), national exams at age 16 (O levels/GCSEs/CSEs or equivalent), optional national exams at ages 17-18 (A levels/AS levels or equivalent), and college or university degree. Smoking status was categorised as never, previous, and current. Employment status was classified as paid or self-employment, retired, looking after home and/or family, unemployed, doing unpaid/voluntary work, unable to work due to sickness or disability, and student. Alcohol intake was categorised as never, special occasions only, 1-3 times a month, 1-2 times a week, 3-4 times a week, and daily or almost daily. For postmenopausal women, all analyses were additionally adjusted for hormone replacement therapy use (never, previous,

current). Missing data was grouped in a separate unknown category for each covariate. There was less than 1% missing data for all covariates, with the exception of educational qualifications in postmenopausal women (17%).

5.3 Results

Descriptive statistics for anthropometric, physical activity and sociodemographic variables are shown in Table 5.1 for premenopausal women and Table 5.2 for postmenopausal women. For all women, the distribution of physical activity was right-skewed (Figure 5.2). For premenopausal women, the mean (SD) physical activity was 31.3 (30.8) excess MET-h/week. For postmenopausal women, the mean (SD) physical activity was 34.4 (33.1) excess MET-h/week. Although the mean total amount of physical activity was higher in postmenopausal than in premenopausal women, postmenopausal women performed less vigorous physical activity than the premenopausal women (mean [SD]: 8.9 [15.6] compared to 9.7 [14.8] excess MET-h/week). Postmenopausal women had higher mean values than premenopausal women for all body composition variables, except for weight, which was nearly identical (on average, premenopausal women were 2 cm taller).

Women who had a high level of physical activity were the least likely to possess a college or university degree while women with a moderate level of physical activity were most likely to hold a college or university degree (Tables 5.1 and 5.2). Those with a high level of physical activity were more likely to have manual and/or standing jobs than those with a low or moderate level of physical activity. Table 5.3 compares characteristics of manual and non-manual workers.

Age-adjusted Pearson's partial correlation coefficients for associations between

different measures of body size and composition are shown in Tables 5.4 and 5.5. BMI was highly correlated with body fat mass in both pre and postmenopausal women (0.94 and 0.93, respectively). BMI showed only a moderate correlation with WHR (0.46 and 0.44 in pre- and postmenopausal women, respectively). WHR was strongly correlated with waist circumference (0.75 in both) but weakly correlated with hip circumference (0.24 and 0.21 in pre- and postmenopausal women, respectively). Total and central adiposity were very highly correlated, as illustrated by the correlation coefficients of body with trunk fat mass (0.98 in both) and body with trunk fat percentage (0.98 in both).

Values for all body size and composition variables were lower with higher amounts of physical activity in both pre-and postmenopausal women (Figure 5.3). The mean BMI for premenopausal women who did less than 1 excess MET-h/week was 27.9 kg/m^2 (95% CI: 27.6, 28.2) compared to 25.0 kg/m^2 (24.9, 25.2) for those who did 75 or more excess MET-h/week of total physical activity. Mean BMI values for postmenopausal women were higher: 28.8 kg/m^2 (28.6, 29.0) for those who did less than 1 MET-h/week and 25.7 kg/m^2 (25.6, 25.8) for those who did 75 or more excess MET-h/week of total physical activity. For premenopausal women, the mean body fat percentage for those who did less than 1 excess MET-h/week was 37.0% (36.5, 37.4%) compared to 31.9% (31.7, 32.2%) for those who did 75 or more excess MET-h/week of total physical activity. Mean body fat percentage was higher for postmenopausal women: 39.4% (39.1, 39.7%) for those who did less than 1 MET-h/week and 34.8% (34.6, 34.9%) for those who did 75 or more excess MET-h/week of total physical activity. Figure 5.3 also shows similar patterns for the other anthropometric variables. The relationships were non-linear, with a greater absolute change in each measure of body size or composition for a

given absolute change in physical activity at lower levels of physical activity. Similar patterns of association between physical activity and adiposity were observed in analysis that were done separately in manual workers and nonmanual workers (Figure 5.4).

Figure 5.5 shows the relationships between BMI, body fat percentage, WHR and total physical activity according to ratios of vigorous-to-total physical activity. Regardless of menopausal status, within each category of total physical activity, an increasing proportion of vigorous physical activity tended to be associated with lower BMI, body fat percentage, and WHR. The proportion of vigorous physical activity had a more distinct association with body composition at higher levels of total physical activity (>30 total excess MET-h/week), where the lines diverge. For premenopausal women, at around 60 total excess MET-h/week, those who did no vigorous physical activity had an average body fat percentage of 33.9% (33.3, 34.6%) compared to 31.6% (31.3, 31.9%) for those whose total physical activity consisted of more than 33% vigorous physical activity. The decrease in body fat percentage from no vigorous physical activity to more than 33% vigorous physical activity was of slightly less magnitude in postmenopausal women: 36.0% (35.8, 36.3%) down to 34.5% (34.3, 34.6%) amongst those who did approximately 60 total excess MET-hours/week of physical activity.

5.4 Discussion

In this large cross-sectional study of 132,793 British women, higher physical activity was associated with a similar reduction in all measures of adiposity for both pre- and postmenopausal women. The majority of the women in our study, espe-

cially at lower levels of physical activity, were overweight, suggesting that it may be more difficult to maintain appropriate energy balance at low levels of physical activity[138]. The relationship between higher physical activity and lower adiposity was steeper for comparisons within the lower range; however, above about 50 excess MET-hours/week, further increase in activity was associated with a relatively small change in measures of adiposity. This may point to a possible considerable benefit of physical activity for those who are most inactive, although it is possible that this pattern of risk may partly reflect reverse causality and/or measurement error, with over-reporting of high physical activity.

The results presented in this chapter agree with previous cross-sectional studies that found a strong correlation between BMI and body fat percentage[139–141]. High age-adjusted correlation coefficients between weight, BMI, waist circumference, hip circumference, body fat percentage, body fat mass, trunk fat percentage, and trunk fat mass suggest that, within our study population, BMI is in fact a fairly good indicator of both total and central adiposity. WHR showed weaker correlation with BMI and the other previously mentioned measurements, with the exception of waist circumference. Accordingly, WHR has shown different associations with disease risk[142] that may capture relevant information beyond adiposity.

There was variation in physical activity by occupation. For instance, 5% of our participants held jobs that usually or always included manual work. Manual workers were characterised by much higher mean levels of both total and vigorous physical activity. They were also less likely to possess a college or university degree than the overall population. The overall means of the various measures of adiposity were similar between those who did and did not hold jobs that at least usually involved manual work. However, within the same level of total physical activity,

manual workers tended to have a higher BMI than non-manual workers. However, the pattern of association between higher physical activity and lower adiposity was similar when comparing manual to non-manual workers.

Higher intensity exercise results in greater energy expenditure for a given duration. After taking into account total energy expenditure, some previous studies suggest that high-intensity exercise is more effective in reducing adiposity than low or moderate-intensity exercise by improving control of energy balance[138,143]. However, other studies have found no additional benefit of higher exercise intensity on abdominal fat loss[144–146] and weight loss[147,148]. Although differential measurement error is possible, our results suggest that a higher proportion of vigorous physical activity within each category of total physical activity was associated with less adiposity. Given the same energy expenditure, high-intensity exercise may stimulate greater fat oxidation and oxygen consumption during the post-exercise period[149,150].

In agreement with previous studies, I found higher values for all body size variables in postmenopausal than in premenopausal women across all categories of physical activity. The menopausal transition results in cessation of ovarian function leading to subsequent decline in oestrogen and progesterone levels, which has been associated with adverse changes in body composition and an increase in central adiposity[151,152]. Studies have reported similar total adiposity and fat distribution in premenopausal and HRT-treated postmenopausal women, suggesting an important role for ovarian hormones in determining fat accumulation and distribution[153,154]. Oestrogen may protect against increased adiposity through regulating appetite, resting metabolic rate, and levels of lipogenesis and lipoprotein lipase activity[155].

A major strength of this study is the availability of a variety of objectively-measured body size and composition measurements. Due to large study size, I was able to exclude women who reported a long-standing illness, disability, or infirmity, which might affect both energy balance and physical activity. I decided to use the IPAQ short form data processing rules due to lack of comprehensive coverage of domain-specific activities on the UK Biobank questionnaire. Although this study is limited by its lack of detailed information on associations between adiposity and levels of specific activities, I felt using questions that captured the frequency and duration of walking, moderate, and vigorous activities was the most straightforward and unbiased approach to estimating physical activity levels.

Limitations include the subjective reporting of physical activity at a single point in time, incomplete reporting of some variables, and lack of data on total energy intake. In this cross-sectional analysis, I cannot assess whether differences in physical activity are causally related to varying levels of adiposity. For instance, those who are highly active may be individuals who are also more successful at maintaining appropriate target energy intake. In Chapter 6, the association between physical activity and adiposity is further explored using UK Biobank accelerometer data, which are objective measures of physical activity. The accelerometer data help to better inform some of the issues around self-reported physical activity. Health outcomes of all UK Biobank participants are being followed long-term, and future analyses will be able to include prospective studies of the relationships between physical activity, body size and composition, and disease outcome[156].

5.5 Conclusion

In conclusion, the analyses presented in this chapter illustrate that UK women reporting more physical activity have lower levels of adiposity, as indicated by a range of body size and composition measures. Additionally, a higher proportion of vigorous physical activity was associated with a further reduction in adiposity. These results suggest that exercise intensity may be associated with body size and composition, beyond the influence of exercise frequency and duration.

Table 5.1: Summary characteristics for the 38,201 premenopausal women participants of UK Biobank by physical activity level

Characteristic	Low PA (<10 MET-h/week)	Moderate PA (10-49.9 MET-h/week)	High PA (≥50 MET-h/week)	All Premenopausal Women
Number of participants	9,142	21,730	7,329	38,201
Age, years: mean (SD)	45.6 (3.5)	45.5 (3.4)	45.6 (3.4)	45.5 (3.4)
Total excess MET-h/week: mean (SD)	5.1 (2.9)	25.3 (10.7)	91.4 (45.9)	31.3 (30.8)
Vigorous excess MET-h/week: mean (SD)	<1 (1.4)	8.4 (8.1)	25.0 (24.7)	9.7 (14.8)
Height, cm: mean (SD)	164.2 (6.3)	164.5 (6.2)	164.3 (6.2)	164.4 (6.2)
Weight, kg: mean (SD)	72.4 (15.2)	69.1 (12.8)	68.0 (12.2)	69.7 (13.4)
Body mass index, kg/m^2: mean (SD)	26.8 (5.4)	25.5 (4.5)	25.2 (4.3)	25.8 (4.8)
< 18.50, %	<1	<1	<1	<1
18.5-24.99, %	42.4	52.8	55.8	50.7
25.00-29.99, %	33.9	31.5	30.9	32.0
≥ 30.00, %	23.0	14.9	12.4	16.3
Body fat percentage: mean (SD)	35.8 (7.0)	33.5 (6.8)	32.4 (6.9)	33.8 (7.0)
Body fat mass, kg: mean (SD)	26.8 (10.7)	23.9 (9.2)	22.7 (8.8)	24.4 (9.6)
Trunk fat percentage: mean (SD)	33.4 (7.9)	31.0 (7.8)	29.6 (8.0)	31.3 (7.9)
Trunk fat mass, kg: mean (SD)	13.6 (5.6)	12.0 (5.0)	11.3 (4.8)	12.3 (5.2)
Waist circumference, cm: mean (SD)	83.3 (12.4)	80.0 (10.8)	79.0 (10.6)	80.6 (11.3)
Hip circumference, cm: mean (SD)	103.6 (10.6)	101.1 (9.1)	100.0 (8.8)	101.5 (9.5)
Waist-to-hip ratio: mean (SD)	0.80 (0.07)	0.79 (0.06)	0.79 (0.06)	0.79 (0.06)
Qualifications [a]				
College or university degree, %	44.1	48.7	39.1	45.7
A levels/AS levels, %	15.8	15.6	15.2	15.6
O levels/GCSEs/CSEs, %	31.5	27.9	34.3	30.0
Vocational qualifications, %	5.8	5.8	7.4	6.1
In paid employment or self-employed, %	91.1	88.3	86.3	88.6
Standing or walking job,% of all participants	15.0	20.5	45.7	24.0
Manual job,% of all participants	2.4	4.6	19.4	7.0
Alcohol consumption				
Never/rarely, %	18.3	15.1	18.8	16.6
Weekly, %	44.4	43.5	44.4	43.9
More than weekly, %	37.3	41.3	36.9	39.5
Smoking status				
Never, %	66.4	65.5	62.9	65.2
Previous, %	23.2	25.9	26.9	25.5
Current, %	10.3	8.4	10.0	9.2
Parity: mean (SD)	1.5 (1.2)	1.5 (1.2)	1.7 (1.2)	1.6 (1.2)
Nulliparous, %	26.8	27.1	24.0	26.4
One or two, %	56.6	54.9	54.1	55.2
Three or more, %	16.6	18.0	21.8	18.4

Abbreviations: Metabolic equivalent hours per week (MET-h/week), physical activity (PA), standard deviation (SD)

[a] Numbers do not add up to 100% due to missing data.

Table 5.2: Summary characteristics for the 94,592 postmenopausal women participants of UK Biobank by physical activity level

Characteristic	Low PA (<10 MET-h/week)	Moderate PA (10-49.9 MET-h/week)	High PA (≥50 MET-h/week)	All Postmenopausal Women
Number of participants	20,756	52,053	21,783	94,592
Age, years: mean (SD)	59.1 (5.4)	59.8 (5.5)	60.3 (5.5)	59.7 (5.5)
Total excess MET-h/week: mean (SD)	5.4 (2.8)	25.5 (10.9)	93.0 (45.9)	34.4 (33.1)
Vigorous excess MET-h/week: mean (SD)	<1 (1.2)	6.4 (7.3)	22.8 (25.7)	8.9 (15.6)
Height, cm: mean (SD)	162.4 (6.2)	162.5 (6.1)	162.0 (6.1)	162.4 (6.1)
Weight, kg: mean (SD)	72.2 (13.6)	69.2 (11.9)	67.9 (11.4)	69.6 (12.2)
Body mass index, kg/m^2: mean (SD)	27.4 (5.0)	26.2 (4.3)	25.9 (4.1)	26.4 (4.5)
< 18.50, %	<1	<1	<1	<1
18.5-24.99, %	34.5	42.7	46.1	41.6
25.00-29.99, %	39.6	39.7	38.4	39.3
≥ 30.00, %	25.4	16.9	14.8	18.2
Body fat percentage: mean (SD)	37.7 (6.4)	36.0 (6.3)	35.1 (6.4)	36.2 (6.4)
Body fat mass, kg: mean (SD)	28.0 (9.6)	25.5 (8.5)	24.4 (8.2)	25.8 (8.8)
Trunk fat percentage: mean (SD)	35.5 (7.2)	33.6 (7.2)	32.5 (7.4)	33.7 (7.3)
Trunk fat mass, kg: mean (SD)	14.3 (5.1)	13.0 (4.6)	12.4 (4.5)	13.1 (4.7)
Waist circumference, cm: mean (SD)	85.7 (12.0)	82.9 (10.9)	81.8 (10.7)	83.2 (11.2)
Hip circumference, cm: mean (SD)	104.2 (10.1)	101.9 (8.9)	100.8 (8.6)	102.1 (9.2)
Waist-to-hip ratio: mean (SD)	0.82 (0.07)	0.81 (0.07)	0.81 (0.07)	0.81 (0.07)
Qualifications [a]				
College or university degree, %	33.1	34.4	27.0	32.4
A levels/AS levels, %	12.2	11.9	10.9	11.7
O levels/GCSEs/CSEs, %	29.1	27.5	28.7	28.1
Vocational qualifications, %	9.9	10.4	12.6	10.8
In paid employment or self-employed, %	59.2	48.0	43.9	49.5
Retired, %	35.7	46.1	49.4	44.6
Standing or walking job,% of all participants	10.7	13.1	25.8	15.5
Manual job,% of all participants	1.3	2.5	9.8	3.9
Alcohol consumption				
Never/rarely, %	20.2	18.3	21.6	19.5
Weekly, %	38.9	37.6	37.3	37.8
More than weekly, %	40.9	44.1	41.0	42.7
Smoking status				
Never, %	58.8	59.7	59.7	59.5
Previous, %	32.8	33.6	33.3	33.4
Current, %	8.2	6.5	7.0	7.0
Parity: mean (SD)	1.8 (1.1)	1.9 (1.1)	1.9 (1.1)	1.9 (1.1)
Nulliparous, %	17.4	17.0	14.5	16.5
One or two, %	59.5	58.9	59.0	59.1
Three or more, %	23.1	24.0	26.4	24.4
Hormone replacement therapy use				
Never, %	54.0	53.1	51.4	52.9
Previous, %	44.0	44.7	46.3	44.9
Current, %	1.9	2.0	2.2	2.0

Abbreviations: Metabolic equivalent hours per week (MET-h/week), physical activity (PA), standard deviation (SD). [a] Numbers do not add up to 100% due to missing data.

Table 5.3: Characteristics of manual and non-manual workers for pre- and post-menopausal women

Characteristic	Premenopausal Women		Postmenopausal Women	
	Manual workers	Non-manual workers	Manual workers	Non-manual workers
Number of participants	2,647	35,554	3,701	90,891
Age, years: mean (SD)	45.3 (3.4)	45.6 (3.4)	56.3 (5.0)	60.0 (5.4)
Total excess MET-h/week: mean (SD)	76.7 (68.3)	30.1 (31.9)	84.1 (74.2)	34.9 (37.4)
Vigorous excess MET-h/week: mean (SD)	17.4 (29.4)	9.2 (14.0)	18.6 (35.0)	8.6 (15.5)
Height, cm: mean (SD)	163.5 (6.3)	164.5 (6.2)	161.8 (6.2)	161.8 (6.2)
Weight, kg: mean (SD)	69.8 (13.6)	69.7 (13.4)	69.6 (13.1)	69.6 (12.2)
Body mass index, kg/m^2: mean (SD)	26.1 (5.0)	25.8 (4.7)	26.6 (4.7)	26.4 (4.4)
Body fat percentage: mean (SD)	33.7 (7.2)	33.8 (6.9)	35.6 (6.7)	36.2 (6.4)
Body fat mass, kg: mean (SD)	24.4 (9.8)	24.4 (9.6)	25.5 (9.3)	25.8 (8.7)
Trunk fat percentage: mean (SD)	30.8 (8.2)	31.3 (7.9)	32.9 (7.7)	33.8 (7.3)
Trunk fat mass, kg: mean (SD)	12.1 (5.2)	12.3 (5.2)	12.8 (5.0)	13.1 (4.7)
Waist circumference, cm: mean (SD)	81.2 (11.8)	80.6 (11.3)	83.4 (11.8)	83.2 (11.1)
Hip circumference, cm: mean (SD)	101.6 (9.9)	101.4 (9.5)	102.0 (9.7)	102.1 (9.1)
Waist-to-hip ratio: mean (SD)	0.80 (0.07)	0.79 (0.06)	0.82 (0.07)	0.81 (0.07)
College or university degree, %	24.9	47.3	17.4	33.0

Table 5.4: Age-adjusted Pearson's partial correlation coefficients for the 38,201 premenopausal women participants of UK Biobank

	Height, cm	Weight, kg	BMI, kg/m^2	Waist, cm	Hip, cm	WHR	Body fat %	Body fat mass, kg	Trunk fat %	Trunk fat mass, kg
Height, cm	1.00									
Weight, kg	0.30	1.00								
BMI, kg/m^2	-0.10	0.92	1.00							
Waist, cm	0.07	0.86	0.87	1.00						
Hip, cm	0.17	0.91	0.88	0.82	1.00					
WHR	-0.08	0.41	0.46	0.75	0.24	1.00				
Body fat %	0.05	0.84	0.86	0.82	0.82	0.45	1.00			
Body fat mass, kg	0.17	0.97	0.94	0.88	0.91	0.44	0.94	1.00		
Trunk fat %	0.16	0.81	0.78	0.77	0.79	0.40	0.98	0.90	1.00	
Trunk fat mass, kg	0.25	0.94	0.87	0.85	0.88	0.42	0.95	0.98	0.95	1.00

Abbreviations: BMI, body mass index; WHR, waist-to-hip-ratio

Table 5.5: Age-adjusted Pearson's partial correlation coefficients for the 94,592 postmenopausal women participants of UK Biobank

	Height, cm	Weight, kg	BMI, kg/m^2	Waist, cm	Hip, cm	WHR	Body fat %	Body fat mass, kg	Trunk fat %	Trunk fat mass, kg
Height, cm	1.00									
Weight, kg	0.31	1.00								
BMI, kg/m^2	-0.12	0.90	1.00							
Waist, cm	0.06	0.85	0.85	1.00						
Hip, cm	0.15	0.89	0.87	0.80	1.00					
WHR	-0.07	0.39	0.44	0.75	0.21	1.00				
Body fat %	0.02	0.82	0.84	0.79	0.80	0.42	1.00			
Body fat mass, kg	0.17	0.96	0.93	0.86	0.90	0.42	0.93	1.00		
Trunk fat %	0.14	0.78	0.75	0.73	0.75	0.38	0.98	0.90	1.00	
Trunk fat mass, kg	0.26	0.93	0.85	0.82	0.86	0.40	0.94	0.98	0.94	1.00

Abbreviations: BMI, body mass index; WHR, waist-to-hip-ratio

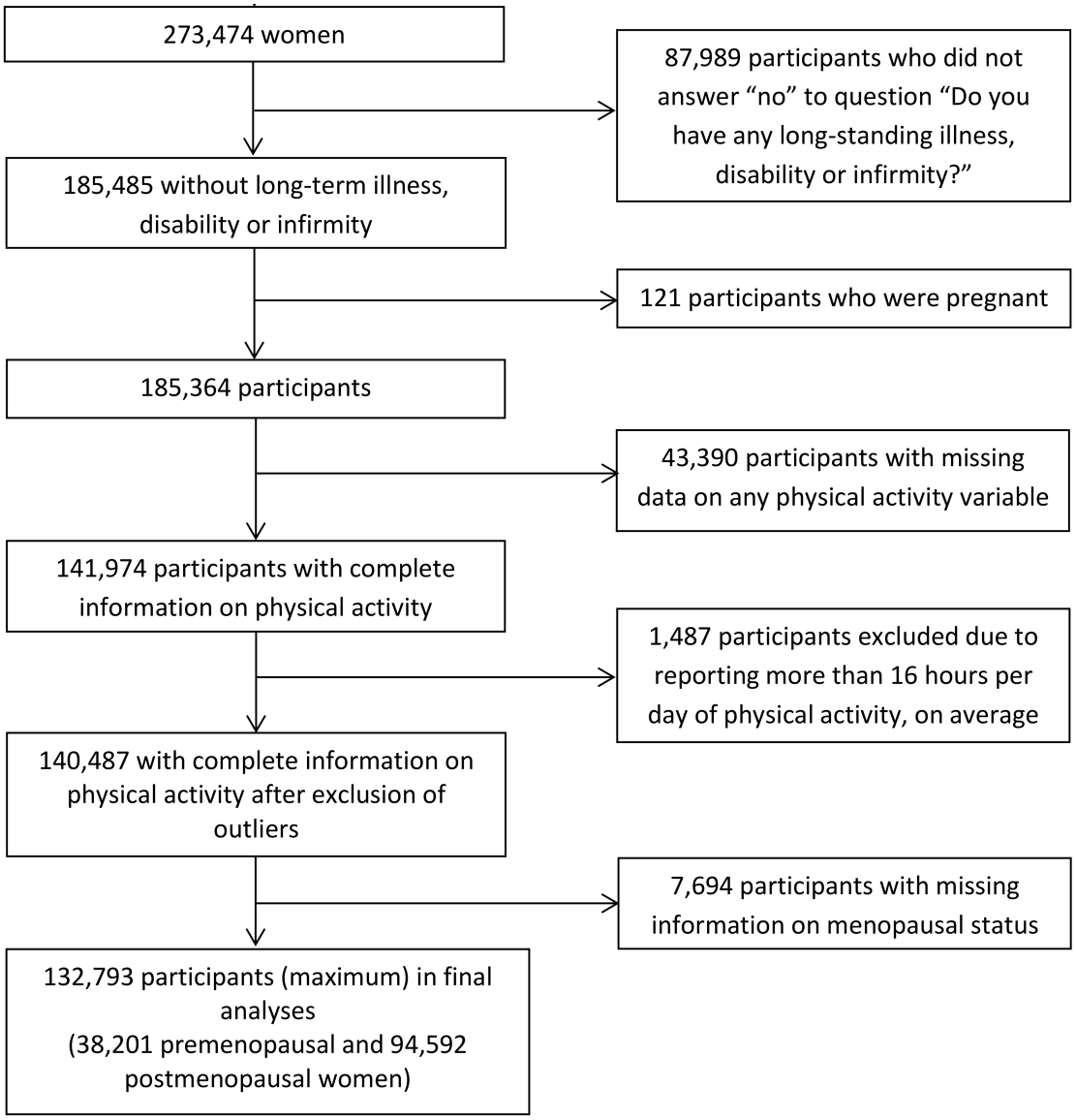
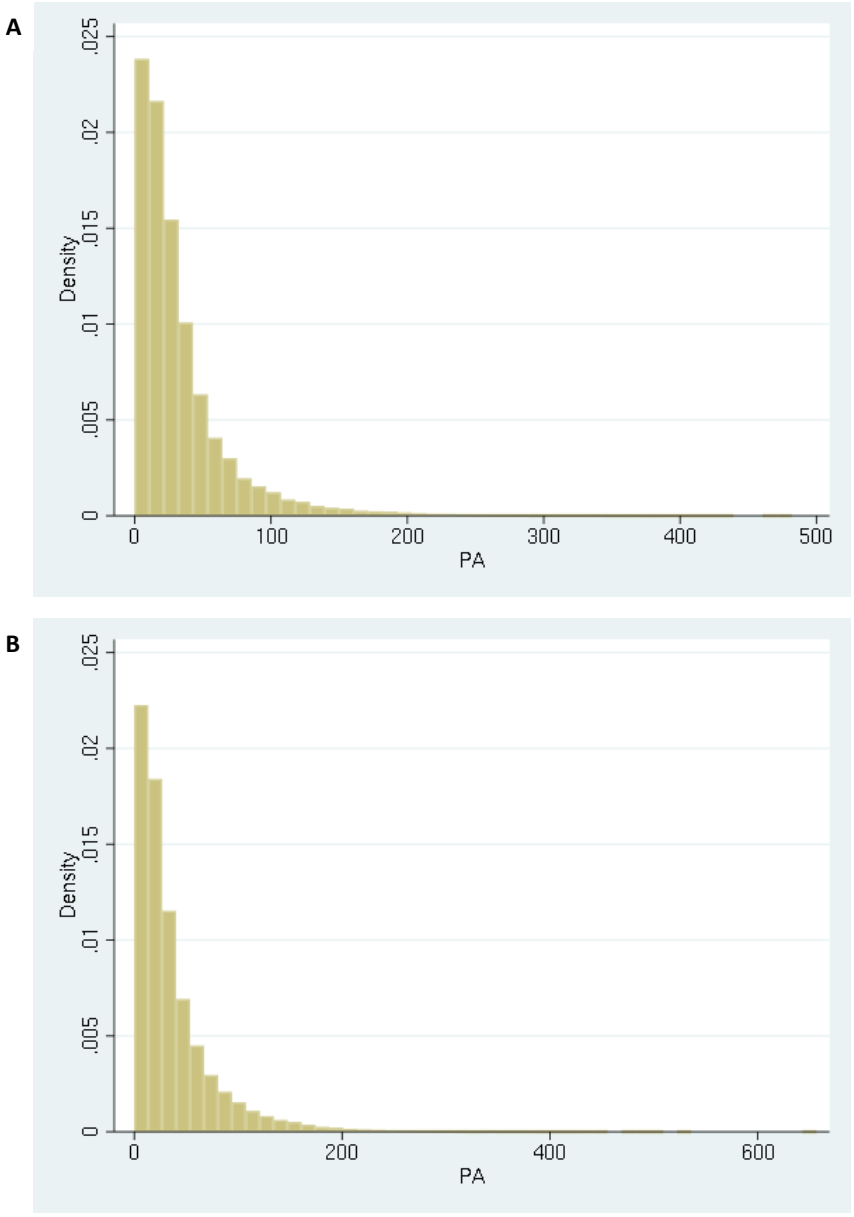
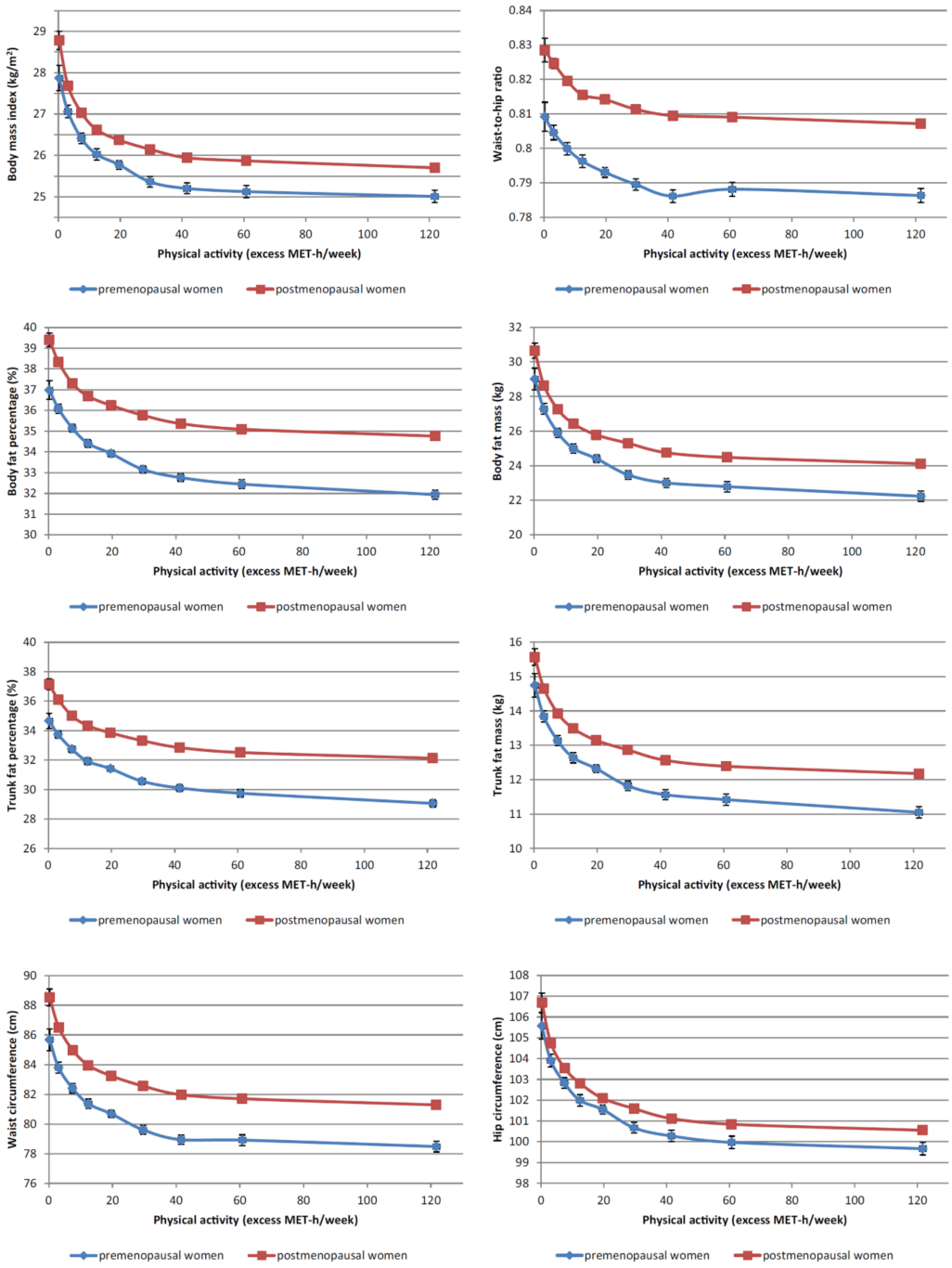


Figure 5.1: Participant inclusion/exclusion flow diagram for analyses of physical activity in relation to body size and composition



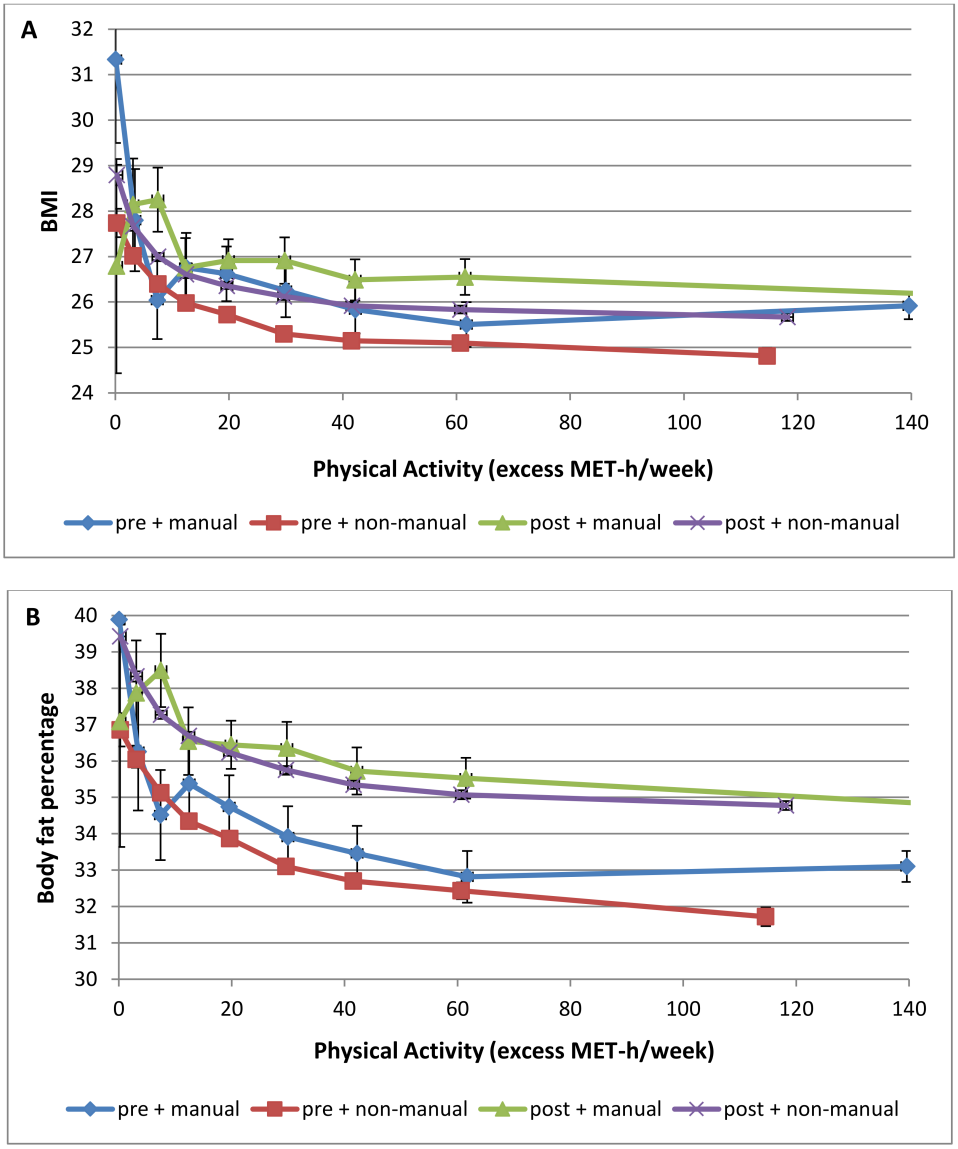
Abbreviation: physical activity (PA)

Figure 5.2: Distribution of physical activity (excess MET-h/week) in premenopausal (A) and postmenopausal (B) women.



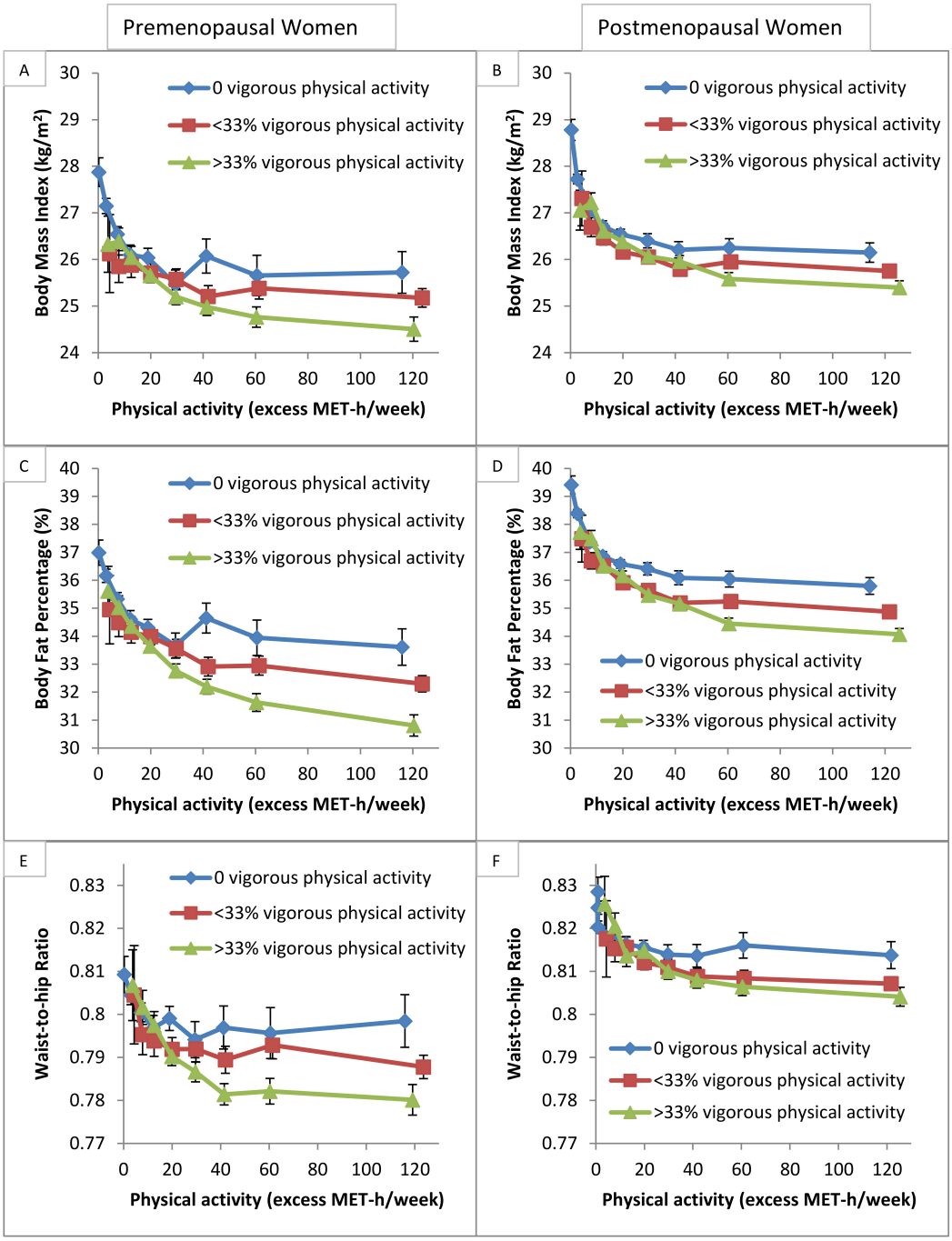
Physical activity was grouped in the following categories: 0, 1-4.9, 5-9.9, 10-14.9, 15-24.9, 25-34.9, 35-49.9, 50-74.9, and ≥75 excess MET-h/week, and are plotted at the value of the mean excess MET-h/week in each category. Error bars are 95% confidence intervals.

Figure 5.3: Association between physical activity and body size/composition by menopausal status for the 38,201 premenopausal and 94,592 postmenopausal women of UK Biobank



Physical activity was grouped in the following categories: 0, 1-4.9, 5-9.9, 10-14.9, 15-24.9, 25-34.9, 35-49.9, 50-74.9, and ≥75 excess MET-h/week, and are plotted at the value of the mean excess MET-h/week in each category.
Error bars are 95% confidence intervals.

Figure 5.4: Association between physical activity and A) BMI and B) Body fat percentage by menopausal status, separately in manual workers and non-manual workers.



Panels A, C, E correspond to premenopausal women and panels B, D, and F correspond to postmenopausal women. Physical activity was grouped in the following categories: 0, 1-4.9, 5-9.9, 10-14.9, 15-24.9, 25-34.9, 35-49.9, 50-74.9, and ≥75 excess MET-h/week, and are plotted at the value of the mean excess MET-h/week in each category. Error bars are 95% confidence intervals.

Figure 5.5: Association between total physical activity and body size/composition by categories of vigorous physical activity for 38,201 premenopausal and 94,592 postmenopausal women.

Chapter 6

ACCELEROMETER AND QUESTIONNAIRE MEASURES OF PHYSICAL ACTIVITY IN RELATION TO BODY SIZE AND COMPOSITION

6.1 Introduction

Previous large-scale observational studies of the association between physical activity and adiposity are mostly based on self-reported physical activity from questionnaires, which is prone to both inaccurate reporting and reporting bias[9]. Prior studies have demonstrated low to moderate correlation between self-reported and objective accelerometer measures of physical activity[157,158]. Self-reported and

accelerometer-measured physical activity capture different aspects of physical activity with limitations unique to each[158]. However, research methods utilising more objective measures of physical activity, along with more detailed measures of body fat, are needed to reduce measurement error and more accurately characterise the association between physical activity and adiposity.

The current chapter investigates the association between physical activity and adiposity, with accelerometer-measured compared to self-reported physical activity in nearly 43,000 women. These associations were assessed using various measures of adiposity, including BMI, body fat percentage, and waist circumference measured by both bioelectrical impedance and dual-energy X-ray absorptiometry (DXA). I also explore how the associations vary by age.

6.2 Methods

6.2.1 Data source

Details of UK Biobank design, rationale, and survey methods have been described elsewhere[110].

6.2.2 Study participants

At the time analyses were run for this chapter, the complete UK Biobank dataset included 502,617 UK adults (229,164 men and 273,453 women) between 40 to 69 years of age at recruitment during 2006 to 2010. The present study was restricted to participants with available accelerometer data (n=103,705). Participants were excluded if they did not have at least 72 hours of data and also data in each

one-hour period of the 24-hour cycle across multiple days (n=6,995). Participants were also excluded if they had insufficient data for calibration (n=4). Participants who had missing data on any of the physical activity variables used in the analyses were excluded (n=15,999). Participants who reported physical activity greater than an average of 16 hours per day (n=620) were also excluded as recommended by the International Physical Activity Questionnaire (IPAQ) scoring guidelines (Appendix B). Participants with missing data on BMI (n=146), body fat percentage (n=988), and waist circumference (n=6) were excluded. Finally, men were excluded (n=35,955) and the final analyses included 42,992 women (Figure 6.1).

6.2.3 Self-reported physical activity

Physical activity questions from the baseline questionnaire captured the frequency and duration of three intensities of activity (walking, moderate, and vigorous). Participants were asked how many days per week they typically engaged in each category of activity. For each category in which an answer of one or more days was given, the participant was subsequently asked the number of minutes on average spent on the activity per day. Questions were adapted from the IPAQ, a validated survey instrument[135], and are listed in Appendix A. Metabolic equivalents (METs) were used to quantify physical activity; 1 MET is expended by sitting quietly for 1 hour, and the MET value reflects the ratio of energy expended per kilogram of body weight per hour to that expended when sitting quietly[136]. The number of minutes per day engaged in each level of activity was multiplied by the respective MET score for the corresponding level of activity (3.3 for walking, 4.0 for moderate physical activity, and 8.0 for vigorous physical activity). MET minutes per day

were converted to MET hours per week. The total amount of METs was calculated by summing total METs from the walking, moderate, and vigorous activity levels. Following IPAQ scoring guidelines, physical activity of less than 10 minutes per day for any category was recoded to 0. This was done in order to stay consistent with other studies that have used IPAQ guidelines. When the IPAQ guidelines were formulated, the scientific evidence suggested that episodes of at least 10 minutes of physical activity were required to achieve health benefits, as stated directly in the guidelines. In more recent years, in some cases after the analyses presented in this thesis were conducted, this notion is being re-evaluated as more studies are conducted.

6.2.4 Accelerometer-measured physical activity

Accelerometer data was available on 58,298 women in the data used for the analyses in this chapter. Participants, all of whom had provided a valid email address, were invited to participate in a seven day accelerometer study between February 2013 and December 2015 (baseline recruitment for the cohort occurred between 2006-2010). Starting in June 2013, participants were sent wrist-worn triaxial accelerometers (Axivity AX3, Newcastle upon Tyne, UK) that were programmed to capture three-dimensional acceleration data at 100 Hz with a dynamic range of ± 8 g. Participants were also given instructions to wear the accelerometer on their dominant wrist continuously for seven days and then to send the device to the coordinating centre using the provided prepaid envelope. The accelerometer variable used in the current analyses represents overall physical activity and is the average vector magnitude of acceleration. Further details on data collection, processing,

and analysis can be found elsewhere[111].

6.2.5 Anthropometry and body composition

Details on anthropometric and body composition measures are provided in Chapter 5. DXA was used to measure fat percentage on a subset of participants beginning in 2014 using the GE-Lunar iDXA (GE Healthcare, Chicago, USA).

6.2.6 Statistical analyses

Baseline characteristics were summarised by physical activity (least active fifth, most active fifth, and overall). Since self-reported physical activity was not normally distributed, Spearman's correlation coefficients were used to measure the strength of correlations between self-reported and accelerometer-measured physical activity in the overall population and in subgroups based on sociodemographic characteristics. Self-reported and accelerometer-measured physical activity were categorised into tenths and the median value within each category of physical activity is shown in the figures. The associations of physical activity and adiposity measures were examined using multivariable linear regression. Analyses comparing the association of accelerometer-measured physical activity with body fat percentage, measured by bioelectrical impedance and DXA were restricted to participants with both measures. Likelihood ratio tests were used to assess whether the associations between physical activity (comparing self-reported and accelerometer-measured) and adiposity were modified by age ($<$55 years or $\geq$55 years).

Covariates were determined a priori and were 5-year age at recruitment cat-

egories, socioeconomic status as indicated by fifths of Townsend deprivation index[119], educational qualifications, employment status, smoking status (never, previous, current), alcohol intake frequency, parity (nulliparous, 1, 2, 3, 4 or more births), and hormone replacement therapy use (never, previous, current). As a covariate, educational qualification was grouped into the following categories: vocational qualifications, national exams at age 16 (O levels, GCSEs, CSEs, or equivalent), optional national exams at ages 17-18 years (A levels, AS levels, or equivalent), and college or university degree. Employment status was categorised as paid or self-employed, retired, looking after home and/or family, unemployed, doing unpaid or voluntary work, unable to work due to sickness or disability, and student. Alcohol intake was categorised as never, special occasions only, 1-3 times a month, 1-2 times a week, 3-4 times a week, and daily or almost daily.

Missing data were grouped in a separate unknown category for each covariate. There were less than 1% missing data for all covariates except for educational qualifications (7.4% missing data). To assess the impact of missing values, a sensitivity analysis restricted to participants with known values for all covariates was conducted. We also conducted sensitivity analyses to assess the impact of excluding participants who reported long-term illness, disability or infirmity, participants who reported fair or poor health rather than excellent or good health, and participants whose jobs usually or always required heavy manual work. Analyses were conducted using STATA, version 15.0 (Stata Corp LP, College Station, TX).

6.3 Results

Characteristics of the study population by least active and most active fifth of accelerometer-measured physical activity are shown in Table 6.1. Mean physical activity measured by accelerometer was 28.7 (SD 8.0) milli-gravity. The most active participants were on average younger and had lower values for all body size and composition variables. They were more likely to have a college or university degree, be employed rather than retired, have an active job, and consume alcohol at least weekly. The least active participants were more likely to be ever smokers and were also more likely to have a long-standing illness or disability.

The correlation between questionnaire and accelerometer-measured physical activity was 0.22 (95% CI: 0.21, 0.23) (Table 6.2). The correlations were comparatively higher in participants who were younger and in participants who had lower BMI. The correlations were lower among participants who reported that their job usually or always involved heavy manual work and/or mainly walking or standing.

The inverse associations between physical activity and all measures of adiposity were linear and approximately twofold larger in models that used accelerometer-measured rather than self-reported physical activity. Comparing the top to bottom tenth of accelerometer-measured physical activity, the difference in BMI was 4.8 (95% CI: 4.6, 5.0) kg/m^2 (Figure 6.2, Table 6.3). Women in the top tenth of accelerometer-measured physical activity had an 8.1% (95% CI: 7.8, 8.3) lower body fat percentage while women in the top tenth of self-reported physical activity had a 4.3% (95% CI: 4.0, 4.5) lower body fat percentage, compared to those in the bottom tenth of physical activity (Figure 6.2, Table 6.3). There was also a nearly twofold larger inverse association between waist circumference and physical

activity when measured by accelerometer rather than questionnaire (Figure 6.2, Table 6.3). The results of sensitivity analyses excluding participants who had any missing values, reported a long-term illness or disability, reported a health rating worse than "good", or whose jobs usually or always required heavy manual work did not materially differ from the main findings.

Table 6.4 shows the associations between accelerometer-measured physical activity and bioelectrical impedance-measured body fat percentage at baseline (2006-2010) compared to body fat percentage measured by DXA starting in May 2014. Body fat percentage by impedance at baseline was lower than body fat percentage by DXA, measured on average seven years later. For both measures, there was a linear dose-response association between physical activity and body fat percentage (Figure 6.3). The inverse associations were stronger when body fat percentage was measured by DXA. Compared to the least active women, the most active women had an 8.8% (95% CI: 7.7, 10.0) lower DXA-measured body fat percentage and a 7.0% (95% CI: 5.9, 8.1) lower impedance-measured body fat percentage (Table 6.4).

Associations between physical activity and measures of adiposity by age group are shown in Figure 6.4. For a given level of accelerometer-measured physical activity, the older participants (over age 55) had a slightly lower BMI but a higher body fat percentage compared to their younger counterparts. There was heterogeneity by age in the association between self-reported physical activity and body fat percentage (P=0.03) but there was no heterogeneity by age when physical activity was measured by accelerometer (P=0.27).

6.4 Discussion

In this large cross-sectional study of nearly 43,000 participants, associations between physical activity and BMI, body fat percentage, and waist circumference were stronger when physical activity was measured by accelerometer compared to questionnaire self-reports. Body fat percentage measured by DXA at follow-up showed a slightly stronger association with physical activity compared to body fat percentage measured by bioelectrical impedance at baseline, but the overall pattern of association was similar. The correlation between accelerometer-measured and self-reported physical activity was lower in participants with higher BMI and in older participants.

The analyses based on accelerometer-measured physical activity suggest an approximately linear inverse association between physical activity and adiposity, with the most active participants having the lowest BMI, body fat percentage, and waist circumference. In contrast, the analyses in the same participants based on self-reported physical activity suggest a comparatively small further benefit of physical activity greater than 50 MET-hours a week on adiposity.

The results in Chapter 5 suggested that the steeper relationship between physical activity and lower adiposity within the lower range of physical activity could be due to either a comparatively larger benefit of physical activity for those who are relatively inactive or measurement error from over-reporting of high physical activity. The present analyses demonstrating an approximately linear dose-response relationship between accelerometer-measured physical activity and adiposity supports the latter explanation and further suggests that over-reporting of total physical activity contributed to the low overall correlation between self-reported and

accelerometer-measured physical activity. Although wrist accelerometer-measured physical activity has limitations, such as measuring movement of only one part of the body and the inability to reliably capture activities such as cycling[158], it has the major advantage of eliminating both inaccurate reporting that leads to random error as well as reporting bias that may vary by sociodemographic characteristics.

Measurement error in the self-reported data results in misclassification of individuals by physical activity status. I used the IPAQ short form data processing rules since the UK Biobank questionnaire did not comprehensively cover domain-specific activities, but it is still likely that lower intensity activities were underreported and reported less accurately[159]. In contrast, the accelerometers were worn for 24 hours a day, over 7 days. Therefore, the lower correlation between self-reported and accelerometer-measured physical activity in older participants and the heterogeneity by age seen only with the self-reported data may be explained by the observation that, in older adults, a greater proportion of physical activity is of lower intensity[160].

Individuals with higher body fat percentage may report moderate and strenuous physical activity less accurately than leaner individuals, based on comparisons between self-reported physical activity and energy expenditure estimated from whole-room indirect calorimeter[161]. In agreement with some previous studies, I found that the correlation between physical activity measured by questionnaire and accelerometer-measured physical activity was greater for those with lower BMI[158]. This suggests that measurement error of self-reported physical activity may be greater in overweight and obese BMI groups.

To my knowledge, the present study is the largest to date comparing accelerometer-measured and self-reported physical activity in relation to direct measures of body

fat, although the results are consistent with prior, smaller studies that suggest a stronger association between adiposity and accelerometer-measured compared to self-reported physical activity[162–166]. A major strength of this study is the availability of both accelerometer-measured physical activity and body fat by impedance in nearly 43,000 participants, together with data on body fat assessed by DXA in over 1,200 women. Additionally, the accelerometers used in this study were waterproof[111], overcoming a limitation of prior studies where the devices had to be removed for water-based activities[166].

While self-reported physical activity was available at baseline in these data, accelerometer-measured physical activity was assessed only 3-5 years after end of recruitment, which raises the question of whether higher adiposity at baseline predicts lower physical activity levels[167] rather than physical activity determining adiposity. However, the analysis of accelerometer-measured physical activity in relation to DXA-measured body fat percentage, which was assessed within the same time frame as accelerometer-measured physical activity, showed similar results to the main analysis based on body fat percentage assessed by impedance at baseline. Other limitations include the lack of data on total energy intake for the whole cohort. Due to the cross-sectional nature of this study, I cannot assess to what extent physical activity is causally related to adiposity. Highly active individuals may also be more likely to maintain appropriate target energy intake, for example.

Future research could incorporate the accelerometer data to calibrate the self-reported measures of physical activity data. Additionally, data is available on acceleration intensity distribution. In the future, it may be possible to use a certain intensity threshold cutoff to differentiate between moderate-to-vigorous intensity activity compared to lower intensity activity. These intensity distribution

variables were not used in the present analyses, because further research needs to be conducted in order to determine and validate any intensity threshold cutoff. Furthermore, Willets et al. have recently developed physical activity phenotypes using a machine learning model with reference behaviors provided by data from a subset of participants who wore a camera along with the accelerometer. These physical activity phenotypes will be incorporated into UK Biobank in the future and will improve the interpretability of the accelerometer data[168]

6.5 Conclusion

In conclusion, the findings based on objective accelerometer data indicate a stronger relationship between physical activity and adiposity than previously thought. Comparison of results with physical activity measured by questionnaire and accelerometer suggest substantial measurement error in self-reported physical activity, emphasising the need to incorporate objective measures of physical activity in future studies.

Table 6.1: Baseline characteristics according to fifths of accelerometer-measured physical activity

	Least active women	Most active women	All women
	<22.2 milli-gravity	**>=34.6 milli-gravity**	
Number of participants	8,606	8,595	42,992
Age at recruitment (years), mean (SD)	58.0 (7.4)	52.6 (7.4)	55.3 (7.7)
Lowest fifth of socioeconomic status	1,897 (22.0%)	1,699 (19.8%)	8,744 (20.3%)
Accelerometer-measured physical activity (milli-gravity), mean (SD)	18.9 (2.7)	40.6 (6.3)	28.7 (8.0)
Self-reported physical activity (MET-hours/week), median (IQR)	21.3 (9.9, 41.7)	40.2 (21.8, 73.2)	29.3 (14.4, 55.3)
Height (cm), mean (SD)	163.2 (6.3)	163.7 (6.1)	163.5 (6.2)
Weight (kg), mean (SD)	75.5 (15.6)	65.0 (10.3)	69.9 (13.2)
BMI (kg/m^2), mean (SD)	28.3 (5.7)	24.3 (3.7)	26.2 (4.8)
Body fat percentage (%), mean (SD)	38.7 (6.6)	31.7 (6.4)	35.3 (6.8)
Waist circumference (cm), mean (SD)	87.9 (13.3)	77.6 (9.5)	82.4 (11.7)
College or university degree	3,586 (41.7%)	4,060 (47.2%)	19,214 (44.7%)
Current employment status			
Paid employment/self-employed	4,401 (51.1%)	6,101 (71.0%)	26,693 (62.1%)
Retired	3,517 (40.9%)	1,591 (18.5%)	12,710 (29.6%)
Other	688 (8.0%)	903 (10.5%)	3,589 (8.3%)
Job involves mainly walking/standing [a]	893 (20.3%)	1,926 (31.6%)	6,648 (24.9%)
Job involves heavy manual work [b]	170 (3.9%)	576 (9.4%)	1,567 (5.9%)
Weekly or more frequent alcohol intake	5,295 (61.5%)	6,292 (73.2%)	29,829 (69.4%)
Ever smoker	3,583 (41.6%)	3,212 (37.4%)	16,936 (39.4%)
Long-standing illness or disability	3,145 (36.5%)	1,449 (16.9%)	1,0685 (24.9%)

Abbreviations: interquartile range (IQR), metabolic equivalent (MET), standard deviation (SD)

[a] Participants who reported their work "usually" or "always" involved walking or standing for most of the time

[b] Participants who reported their work "usually" or "always" involved heavy manual or physical work for most of the time

Table 6.2: Spearman correlation between self-reported physical activity (in MET-hours) and accelerometer-measured physical activity (in milli-gravity), according to participant characteristics

	N Women (%)	Correlation	95% Confidence Interval
Total	42,992	0.22	0.21, 0.23
Age group at recruitment (years)			
<55 years	18,973 (44.1)	0.26	0.25, 0.28
55+ years	24,019 (55.9)	0.20	0.19, 0.22
Socioeconomic status, fifths			
Top fifth	8,401 (19.5)	0.22	0.30, 0.24
Bottom fifth	8,744 (20.3)	0.22	0.30, 0.24
BMI (kg/m^2)			
<25	20,255 (47.1)	0.21	0.20, 0.23
25-29.9	15,146 (35.2)	0.18	0.17, 0.20
≥30	7,591 (17.7)	0.15	0.13, 0.17
College or university degree			
Yes	19,214 (44.7)	0.22	0.21, 0.24
No	23,778 (55.3)	0.22	0.20, 0.23
Current employment status			
In paid employment or self-employed	26,693 (62.1)	0.24	0.23, 0.25
Retired	12,710 (29.6)	0.22	0.20, 0.24
Other	3,589 (8.4)	0.30	0.27, 0.33
Job involves mainly walking/standing			
Never or rarely	12,191 (45.7)	0.25	0.23, 0.27
Sometimes	7,839 (29.4)	0.21	0.19, 0.23
Usually or Always	6,648 (24.9)	0.18	0.16, 0.20
Job involves heavy manual work			
Never, rarely	20,762 (77.8)	0.24	0.22, 0.25
Sometimes	4,353 (16.3)	0.17	0.14, 0.20
Usually or Always	1,567 (5.9)	0.13	0.08, 0.18
Alcohol intake frequency			
Weekly or more	29,829 (69.4)	0.22	0.21, 0.23
Less than weekly	13,152 (30.6)	0.21	0.20, 0.23
Smoking status			
Never	25,998 (60.5)	0.21	0.20, 0.22
Ever	16,936 (39.4)	0.23	0.22, 0.25
Long-standing illness or disability			
No	32,307 (75.2)	0.21	0.20, 0.22
Yes	10,685 (24.9)	0.23	0.21, 0.24

Table 6.3: Association of self-reported questionnaire and accelerometer-measured physical activity with measures of adiposity

	N	BMI (kg/m^2) Difference (95% CI)	Body fat percentage (%) Difference (95% CI)	Waist circumference (cm) Difference (95% CI)
Questionnaire				
Bottom tenth (least active)	4,433	0.00 (reference)	0.00 (reference)	0.00 (reference)
2nd tenth	4,187	-0.81 (-1.00, -0.61)	-1.05 (-1.33, -0.78)	-1.77 (-2.25, -1.29)
3rd tenth	4,278	-1.37 (-1.57, -1.18)	-1.89 (-2.16, -1.61)	-3.33 (-3.81, -2.86)
4th tenth	4,318	-1.44 (-1.64, -1.25)	-2.16 (-2.43, -1.88)	-3.74 (-4.22, -3.27)
5th tenth	4,296	-1.60 (-1.79, -1.40)	-2.44 (-2.72, -2.17)	-3.92 (-4.39, -3.44)
6th tenth	4,308	-1.86 (-2.05, -1.66)	-2.89 (-3.16, -2.61)	-4.71 (-5.18, -4.23)
7th tenth	4,276	-1.98 (-2.18, -1.79)	-3.18 (-3.45, -2.90)	-5.02 (-5.50, -4.55)
8th tenth	4,300	-2.20 (-2.39, -2.00)	-3.57 (-3.85, -3.30)	-5.73 (-6.20, -5.25)
9th tenth	4,305	-2.28 (-2.48, -2.09)	-3.86 (-4.14, -3.59)	-5.91 (-6.39, -5.44)
Top tenth (most active)	4,291	-2.51 (-2.70, -2.31)	-4.25 (-4.52, -3.97)	-6.39 (-6.87, -5.91)
Accelerometer				
Bottom tenth (least active)	4,314	0.00 (reference)	0.00 (reference)	0.00 (reference)
2nd tenth	4,292	-1.14 (-1.33, -0.95)	-1.43 (-1.70, -1.16)	-3.02 (-3.49, -2.56)
3rd tenth	4,293	-1.84 (-2.03, -1.65)	-2.35 (-2.62, -2.09)	-4.42 (-4.89, -3.95)
4th tenth	4,307	-2.31 (-2.50, -2.12)	-2.90 (-3.17-2.63)	-5.43 (-5.90, -4.97)
5th tenth	4,312	-2.60 (-2.79, -2.41)	-3.52 (-3.79, -3.25)	-6.27 (-6.74, -5.80)
6th tenth	4,286	-2.92 (-3.11, -2.72)	-4.05 (-4.32, -3.79)	-7.18 (-7.65, -6.71)
7th tenth	4,292	-3.22 (-3.42, -3.04)	-4.63 (-4.90, -4.36)	-7.88 (-8.35, -7.41)
8th tenth	4,301	-3.60 (-3.80, -3.41)	-5.34 (-5.61, -5.07)	-8.80 (-9.27, -8.33)
9th tenth	4,305	-4.17 (-4.36, -3.98)	-6.37 (-6.64,-6.10)	-10.12 (-10.59, -9.64)
Top tenth (most active)	4,290	-4.80 (-4.99, -4.60)	-8.06 (-8.33, -7.78)	-11.92 (-12.39, -11.44)

Abbreviations: body mass index (BMI), confidence interval (CI)
Analyses are adjusted for age, socioeconomic status, alcohol intake, smoking status, educational qualifications, and employment status, parity, and hormone replacement therapy use.

Table 6.4: Association of accelerometer-measured physical activity with body fat percentage, measured by bioelectrical impedance and dual energy x-ray absorptiometry (DXA)

Accelerometer-measured physical activity	N	Mean body fat percentage (95% CI)	Difference (95% CI)
		Measured by impedance	
1st fifth (least active)	270	38.4 (37.6, 39.1)	0.00 (reference group)
2nd fifth	244	36.6 (35.9, 37.4)	-1.46 (-2.55, -0.37)
3rd fifth	265	36.1 (35.4, 36.9)	-1.93 (-3.01, -0.85)
4th fifth	228	34.5 (33.7, 35.3)	-3.43 (-4.55, -2.31)
5th fifth (most active)	265	30.8 (30.1, 31.6)	-6.97 (-8.07, -5.87)
		Measured by DXA	
1st fifth (least active)	270	42.5 (41.7, 43.3)	0.00 (reference group)
2nd fifth	244	40.4 (39.6, 41.2)	-1.87 (-3.04, -0.71)
3rd fifth	265	39.8 (39.0, 40.6)	-2.55 (-3.70, -1.40)
4th fifth	228	37.1 (36.2, 37.9)	-4.98 (-6.17, -3.78)
5th fifth (most active)	265	33.3 (32.5, 34.1)	-8.83 (-10.0, -7.65)

Abbreviations: confidence interval (CI), dual energy x-ray absorptiometry (DXA)
Analyses are adjusted for age, socioeconomic status, alcohol intake, smoking status, educational qualifications, employment status, parity and hormone replacement therapy use.

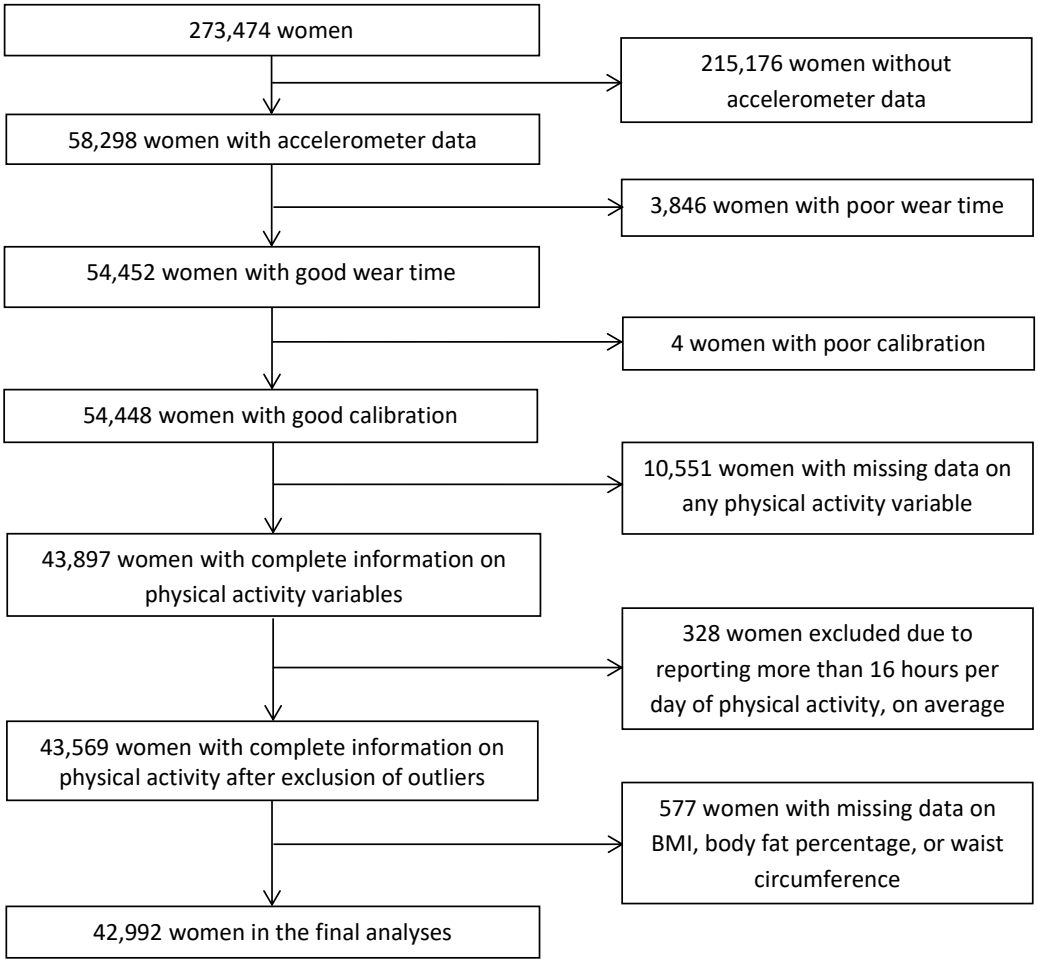
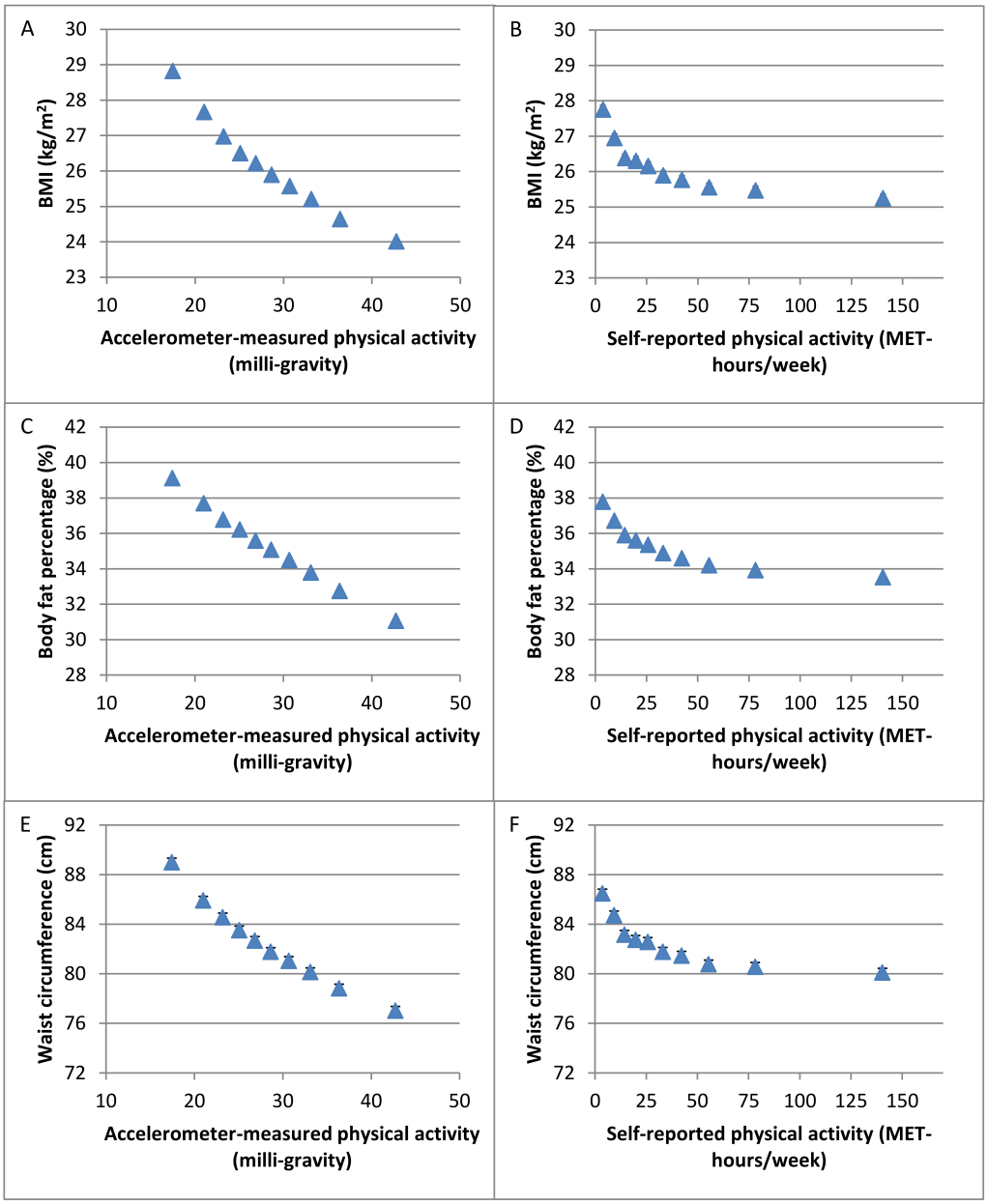
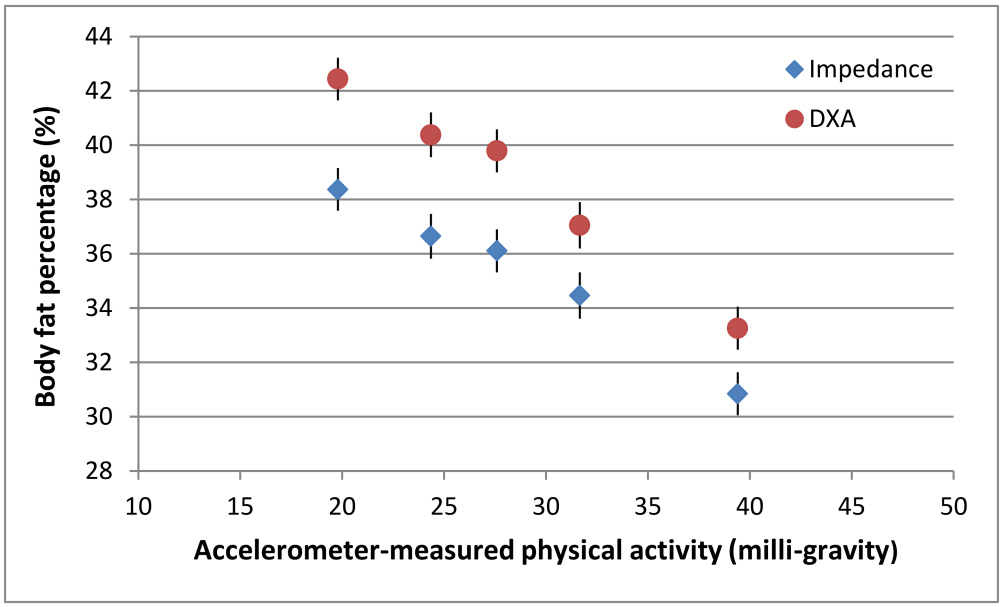


Figure 6.1: Participant inclusion/exclusion flow diagram for analyses of accelerometer-measured and self-reported physical activity in relation to adiposity



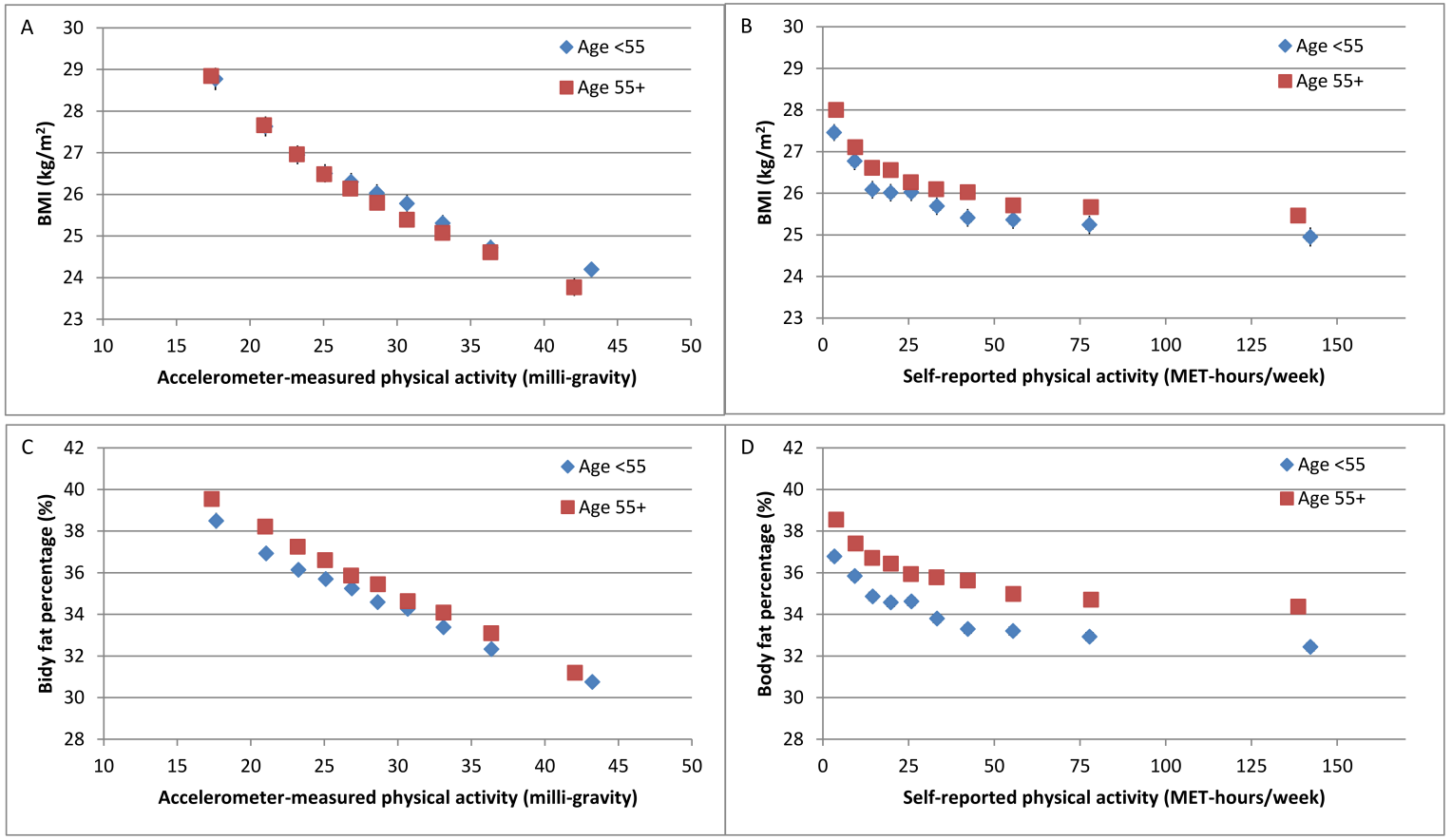
Association of A) accelerometer-measured and B) self-reported physical activity with BMI. Association of C) accelerometer-measured and D) self-reported physical activity with body fat percent. Association of E) accelerometer-measured and F) self-reported physical activity with waist circumference.
Adjusted geometric means (from linear regression models) for BMI, body fat percentage, and waist circumference are plotted against the median value within each tenth of self-reported or accelerometer-measured physical activity. The figure shows point estimates and 95% confidence intervals. Abbreviations: BMI (body mass index), MET (metabolic equivalent)

Figure 6.2: Association of accelerometer-measured and self-reported physical activity with adiposity variables



Physical activity was grouped into fifths. Adjusted geometric means from linear regression models for body fat percentage are plotted against the median value within each fifth of accelerometer-measured physical activity

Figure 6.3: Association of accelerometer-measured physical activity with body fat percentage measured by impedance and dual-energy X-ray-absorptiometry (DXA), 1,272 women included in analysis



Association of physical activity measured by A) accelerometer and B) self-reported questionnaire with BMI. Association of physical activity measured by C) accelerometer and D) self-reported questionnaire with body fat percentage. Adjusted geometric means (from linear regression models) for BMI and body fat percentage are plotted against the median value within each tenth of self-reported or accelerometer-measured physical activity.
The figure shows point estimates and 95% confidence intervals. Abbreviation: metabolic equivalent (MET).

Figure 6.4: Association of self-reported and accelerometer-measured physical activity with adiposity variables by age group

Chapter 7

GENE BY ENVIRONMENT INTERACTIONS IN ADIPOSITY

7.1 Introduction

Individual susceptibility to obesity is determined by the complex interplay of both genetic and environmental factors[169]. Recent genome-wide association studies have identified numerous BMI associated loci[170,171].

As a component of total energy expenditure, physical activity is an important factor in body weight regulation[172]. Various studies have demonstrated that the effect of the FTO variant on adiposity is attenuated among participants who are more physically active[173–184]. Other studies have created genetic risk scores and found greater physical activity may attenuate the genetic predisposition to high BMI[185–187]. Results are, however, inconsistent and a large meta-analysis of 111,421 individuals, comprised of 11 cohorts weighted by study size, found mixed results with a statistically significant interaction effect in North American cohorts but not

in European cohorts[188].

Some studies suggest that genetic predisposition to adiposity may be amplified by greater sedentary behaviour. One study found that a sedentary lifestyle, as measured by television watching, enhances the genetic predisposition to high BMI[189]. Another study reported that the association between two BMI-associated single nucleotide polymorphisms (SNPs) and BMI was modestly strengthened among adolescents with more television, video, and computer game use[190]. In a more recent study, there was a significant interaction of FTO and time spent sitting on BMI[191].

In this chapter, I sought to clarify inconsistencies in the literature by investigating whether the effect of genetic susceptibility, as measured by a comprehensive genetic risk score, on objectively measured BMI was modified by physical activity as well as leisure screen time behaviours in genetically white European individuals within one large cohort, UK Biobank.

7.2 Methods

7.2.1 Data source

Details are provided in prior chapters.

7.2.2 Study sample

At the time the analyses for this chapter were run, the complete UK Biobank dataset included 502,656 UK adults (229,182 men and 273,474 women) between the ages of 40-69 at recruitment during 2006-2010. At the time of these analyses,

genetic data were available on 152,732 participants in total. 112,227 of these patients had complete information available on physical activity and screen time behaviour. These analyses were further restricted to those participants reporting excellent, good or fair health (n=107,275) and to participants who are genetically white European (n=92,738). The final analyses included 46,784 women. The participant inclusion/exclusion criteria are shown in Figure 7.1.

7.2.3 Assessment of BMI, physical activity, and screen time behaviour

Details of the assessment of BMI, self-reported physical activity from questionnaires, and accelerometer-measured physical data are provided in Chapter 6. Physical activity levels were categorised as follows: low ($<$10.0 MET-hours per week), moderate (10-49.9 MET-hours per week), and high ($\geq$50.0 MET-hours per week). These categories are based on IPAQ scoring guidelines, shown in Appendix B.

Screen time behaviour per day was used as an estimate of leisure time sedentary activity. Screen time consisted of the sum of television watching and leisure time computer use. Participants were asked: "In a typical day, how many hours do you spend watching TV? (Put 0 if you do not spend any time doing it)" for television watching. Participants were asked "In a typical day, how many hours do you spend using the computer? (Do not include using a computer at work; put 0 if you do not spend any time doing it)" for computer use. Questions from the UK Biobank touchscreen questionnaire relevant to this thesis are listed in Appendix A.

7.2.4 Genotyping

Characteristics of the 70 SNPs are listed in Table 7.1. For specific reasons given below, I selected 70 of 76 single-nucleotide polymorphisms (SNPs) that showed genome-wide significant association with BMI in the Genetic Investigation of Anthropometric Traits (GIANT) consortium[170]. BMI variants were restricted to those that were associated with BMI in the analysis of individuals of European ancestry. I did not include variants that reached genome-wide significance in only one sex or were classified as a secondary signal within the locus. Three variants were excluded based on failure to meet Hardy-Weinberg equilibrium criteria (p<1x10-6: rs17001654, rs9925964, rs2075650). Two variants did not meet criteria of call rate >95% (rs10733682, rs758747), and one was not available in UK Biobank (rs2033529). The 70 variants that passed quality control criteria in the analyses presented here accounted for approximately 1.5% of the variance in BMI in the GIANT consortium.

7.2.5 Genetic risk score (GRS) computation

Each individual variant was recoded as 0, 1, and 2 according to the number of BMI-increasing risk alleles. Each SNP was weighted by its corresponding per allele log relative risk estimate (β coefficient) obtained from the reported GIANT Consortium meta-analysis data [3]. I first created a weighted score using the following formula where β is the log relative risk representing the association between each SNP and BMI:

Weighted score = $\beta_1 \times$ SNP$_1$ + $\beta_2 \times$ SNP$_2$ + ... + $\beta_n \times$ SNP$_n$

Next, I rescaled the weighted score to reflect the number of BMI-increasing alleles using the following equation:

$$\text{Weighted genetic risk score} = (\text{weighted score} \times \text{number of SNPs})/\text{sum of } \beta \text{ coefficients}$$

The weighted genetic risk score was normally distributed.

7.2.6 Statistical analyses

Linear regression models were used to examine the cross-sectional associations of the weighted GRS with BMI, overall and by each of the three physical activity levels and each of the three screen time behaviour levels. Linear regression models were also used to investigate the associations of physical activity and BMI, overall and by thirds of GRS. Logistic regression models were used to cross-sectionally assess the odds of being obese compared to being normal weight, overall and by physical activity levels, screen time behaviour levels, and GRS levels. Tests for interaction were performed using likelihood ratio tests. Analyses did not allow for interaction between SNPs.

I created a variable of fifths of weighted GRS in order to calculate adjusted geometric means for each combination of GRS fifth across the three categories of physical activity, and the three categories of screen time behaviour.

Analyses were adjusted for age, age^2, self-reported health status and region of recruitment. Adding age-squared as a covariate allowed me to account for any potential non-linear relationship between age and the independent variable, genetic risk score. This approach is commonly used in genetic studies and allows for the effect of age to be modeled at differing ages rather than assuming that effect is

linear across all ages. Adjustment for other potential covariates made no material difference to the results. Analysis of screen time behaviour was further adjusted for physical activity. Statistical analyses were performed using STATA version 14.0 (Stata Corp LP, College Station, TX).

7.3 Results

7.3.1 Characteristics of study participants and association of weighted genetic risk score with BMI

Characteristics of the study participants are shown in Table 7.2. The 2,021 women who had repeat assessment of BMI had a slightly lower median BMI compared to the entire population of this study. The median BMI at repeat assessment 2-3 years after baseline measurements remained essentially the same (Table 7.2). The GRS ranged from 20.8 to 41.5 across the fifths, for the main analyses in this chapter (Table 7.3). Mean BMI was associated with fifths of GRS ($p<0.0001$). The mean BMI within individuals in the lowest fifth of GRS was 25.8 (95% CI 25.7-25.9) kg/m^2 compared to 27.3 (95% CI 27.2-27.4) kg/m^2 amongst those in the top fifth GRS. There was no association between GRS and physical activity measured by accelerometer (Table 7.3).

7.3.2 Adjusted geometric means of BMI by genetic risk score and total physical activity

Adjusted geometric means of BMI stratified by fifths of GRS and total physical activity levels demonstrate decreasing BMI across categories of increasing physical

activity and increasing BMI across categories of increasing GRS. The geometric mean BMI was greatest for those in the top fifth of GRS who were inactive (28.6, 95% CI 28.3-28.8) kg/m^2 compared to those in the bottom fifth of GRS who were most active (25.5, 95% CI 25.3-25.6) kg/m^2 (Table 7.4). Figure 7.2 illustrates geometric mean BMI across fifths of GRS for each level of physical activity.

7.3.3 Association of genetic risk score with BMI and odds of being obese, by physical activity level

Table 7.5 illustrates the effect of weighted GRS on BMI overall and by level of physical activity. Each unit increase in the weighted GRS was associated with a 0.218 (standard error [SE] 0.008) kg/m^2 increase in BMI ($p<0.0001$). GRS was not associated with physical activity level ($p=0.491$). The mean objective accelerometer measures of physical activity (milli-gravity) increased in correspondence with self-reported physical activity (MET-hours/week) from the least active to most active group.

In the low physical activity group (<10 MET-hours/week), each unit increase in weighted GRS was associated with a 0.299 (SE 0.023) kg/m^2 increase in BMI ($p<0.0001$) but the effect was less pronounced in the high physical activity group (0.188 [SE 0.014] kg/m^2), who reported more than 50 MET-hours/week. The effect in the moderately active group (10-49.9 MET-hours/week) was similar to that in the high activity group, 0.207 (SE 0.012) kg/m^2. Comparing the top to bottom fifth of GRS, the most inactive participants had a 2.23 (SE 0.192) kg/m^2 higher BMI while the most active participants had only a 1.33 (SE 0.118) kg/m^2 higher BMI. There was interaction between the GRS and physical activity level on BMI

(p = 0.001). Overall, the odds of being obese versus normal weight was 2.22 (95% CI 2.05-2.41) comparing the top to bottom fifth of GRS. The odds ratio was 2.50 (95% CI 2.09-2.99) in the low activity group compared to 2.20 (95% CI 1.86-2.56) in the highly active group (Table 7.5).

7.3.4 Associations of physical activity with BMI by genetic risk score: comparison of self-reported versus accelerometer data

As established in previous chapters, there is an inverse association between physical activity and BMI. Figure 7.3 shows this association by thirds of GRS with physical activity measured by questionnaire (panel A) compared to with physical activity measured by accelerometer (panel B). There was a larger range of BMI, across fifths of physical activity, when physical activity was measured by accelerometer. Mean BMI in the most active individuals with the lowest GRS was 24.0 (95% CI 23.7-24.3) kg/m^2 by accelerometer and 24.9 (95% CI 24.6-25.1) kg/m^2 by questionnaire. Mean BMI in the least active individuals with the highest GRS was 28.3 (95% CI 28.0-28.7) kg/m^2 by accelerometer and 27.6 (95% CI 27.3-27.9) kg/m^2 by questionnaire. There was a greater trend for increasing BMI, across fifths of accelerometer-measured physical activity, for those with high GRS β=-0.975 (SD 0.061) compared to for those with medium GRS β=-0.730 (SD 0.053) (Table 7.6). Patterns were similar with self-reported physical activity. However, the absolute differences in BMI were greater when physical activity was measured by accelerometer compared to questionnaire self-reports, for all levels of GRS (Table 7.6).

7.3.5 Association of genetic risk score with BMI and odds of being obese, by leisure screen time

Leisure screen time per day, after adjustment for physical activity, modified the effect of GRS on BMI (p = 0.016). However, the magnitude of the interaction was more modest, per unit increase in GRS: 0.191 (SE 0.015) kg/m^2 increase in BMI in those with 2 hours or fewer of screen time per day compared to 0.263 (0.017) kg/m^2 in those with more than four hours per day of screen time. The results were very similar with or without adjustment for physical activity (Table 7.7). Figure 7.4 illustrates geometric mean BMI across fifths of GRS for each level of leisure screen time.

7.4 Discussion

In this large cross-sectional study of 46,784 British women, physical activity attenuated the genetic predisposition to high BMI and risk for obesity. Furthermore, after adjustment for physical activity, there remained an interaction between greater leisure screen time and GRS, in which those who were more sedentary experienced a greater effect of GRS on BMI.

Many of the earlier gene by physical activity interaction studies examined the interaction between the FTO locus and physical activity, with conflicting results. Some report that physical activity attenuated the effect of FTO on adiposity among adults[173–184,192]. Other studies do not find evidence or only find minimal evidence that physical activity modifies the effects of FTO on obesity risk[193,194], possibly due to small sample sizes and measurement error in the assessment of physical

activity, leading to inadequate statistical power to detect an association[195].

However, two studies that used accelerometer-measured physical activity have shown substantial attenuation of the increase in BMI among physically active study participants. In a recent study of 1,280 participants, FTO increased BMI by 0.16 kg/m^2 per allele (p=0.388) among active participants, compared 1.06 kg/m^2 per allele (p=0.024) among inactive participants[176]. In a study of 704 participants, utilising accelerometer-measured physical activity yielded similar results showing that a FTO risk allele was associated with a 1.22 (SD 0.38) increase in BMI in the low physical activity group but only a 0.27 (SD 0.31) increase in BMI in the high physical activity group (P for interaction = 0.004)[182].

In the analyses presented in this chapter, I have used a GRS consisting of 70 BMI-associated SNPs[170] that capture our current best understanding of overall genetic predisposition to obesity. Several previous studies agree with the present findings showing greater physical activity attenuated the effect of a GRS on BMI and on the odds of obesity[185–187,189]. In addition, shortly after the completion of the analyses presented in this thesis, two studies were published using UK Biobank data. Tyrrell et al. demonstrated that physical activity attenuated the association of a 69-SNP GRS on BMI in over 109,000 UK Biobank participants[196]. Furthermore, within the same study, a similar result using accelerometer-measured physical activity was reported in a subset of 19,229 participants who had accelerometer data. A subsequent study in UK Biobank by Rask-Anderson et al. similarly reported gene-environment interaction for variables that captured frequencies and durations of different levels of physical activity[197].

However, an earlier meta-analysis found a statistically significant interaction effect among North American cohorts (n=39,810; P for interaction=0.014) but not

among European cohorts (n=71,611; P for interaction=0.275)[188]. Another study among approximately 8,000 Pakistani participants also showed no interaction between a 95-SNP GRS and physical activity on BMI[198]. It is unclear whether there are true differences in these gene-environment interactions by ethnicity. However, it seems perhaps more likely that the inability of previous large studies to detect gene-environment interactions may be due to small sample sizes and between-study heterogeneity due to the combination of several studies that used different ways of measuring the environmental exposures.

The association of physical activity, measured by self-reported questionnaire compared to accelerometer, with BMI suggests that physical activity may be even more beneficial for lower adiposity overall, and especially in the high GRS group, than results using self-reported physical activity would suggest. For example, in the analysis using accelerometer-measured physical activity, the mean BMI for the most active participants in all three GRS groups was within the healthy BMI range of <25 kg/m^2. In contrast, in the analysis using self-reported physical activity, the corresponding mean BMI was in the healthy range only for participants with the low GRS. Furthermore, the linear trends for the three groups of GRS appear to converge to a greater extent with physical activity measured by accelerometer compared to questionnaire self-report, providing stronger evidence than the self-reported data that genetic predisposition to adiposity can be overcome by higher levels of physical activity. The genetic predisposition to adiposity may be both biological and behavioural, for example, related to lower physical activity and higher caloric intake. These results further illustrate the issues around measurement error in studies of gene-environment interactions[199].

I chose to use leisure screen time as the indicator of sedentary activity, because

leisure sedentary activities may be more easily altered by public health efforts in contrast to occupation-related sedentary activity such as driving or office work. Although there was a statistically significant interaction by leisure screen time per day on the effect of GRS on BMI, the differences in BMI unit increase corresponding to each additional unit increase in the GRS were modest. This may partly be due to lack of an accurate, comprehensive measure of sedentary behaviour. Moreover, the interaction with screen time demonstrated in this present study may not be causal but instead due to other factors such as incomplete adjustment for physical activity, diet, frequency of alcohol consumption, mean sleep duration, socioeconomic status, and smoking. A recent study in UK Biobank investigating whether various lifestyle factors modify the effect of FTO on BMI found no interaction between FTO and television watching in a joint model[175], although it is possible that television watching may not be a comprehensive indicator of sedentary time.

Strengths of this study include a large sample size within a single study, which may help to clarify inconsistent results in the literature from smaller studies published prior to the analyses presented in this chapter, which were conducted in summer 2016. In contrast to prior studies that only looked at FTO or genetic risk scores with fewer BMI-associated variants, I aimed to capture a greater proportion of genetic susceptibility to adiposity by using 70 established BMI-associated variants. Genetic risk scores are better powered to detect gene-environment interactions than individual SNPs[200]. Moreover, objective measures of physical activity were available on a subset of the study population, which provided a more precise measure of physical activity. This is important given that random measurement error associated with exposure assessment typically attenuates the estimate of the

true effect size[199].

Limitations of the study include the use of self-reported physical activity and sedentary behaviour in the main analysis, which were assessed at the same time as BMI was measured at recruitment. The limitations of self-reported physical activity, such as inaccurate reporting and reporting bias, have been discussed extensively in Chapter 6. However, incorporation of the objectively-measured accelerometer values in analyses in this chapter directly comparing the associations between physical activity and BMI by GRS strengthens these findings in that they may mitigate some of these biases. Additionally, heteroscedasticity, defined as unequal variance in data, may be an issue. Heteroscedasticity can result in a false positive result for interaction since a greater mean and variance of BMI in the group with higher BMI compared to the groups with lower BMI may result in a greater effect of the BMI-associated genetic variants driven by statistical artefacts[196].

7.5 Conclusion

The results presented in this chapter demonstrate that higher levels of physical activity attenuate the effect of genetic susceptibility to adiposity on high BMI and obesity risk. The results further highlight the importance of using precise and unbiased measures of physical activity to minimise measurement error in studies of gene-environment interactions. A sedentary lifestyle, represented by screen-time behaviour, may play a role in enhancing the genetic susceptibility to adiposity, although this may be reflective of other associated behaviours and merits further research to better understand the relationships between physical activity and

sedentary activity on adiposity and health outcomes.

Table 7.1: Characteristics of the 70 single nucleotide polymorphisms (SNPs) associated with body mass index at genome wide significance, included in the analyses and the 6 that were excluded

SNP	Chr.	Nearest gene	Effect allele	Other allele	Beta	Imputation quality	Included in score?
rs11165643	1	*PTBP2*	T	C	0.022	0.99575	Yes
rs11583200	1	*ELAVL4*	C	T	0.018	0.98728	Yes
rs12401738	1	*FUBP1*	A	G	0.021	0.99528	Yes
rs12566985	1	*FPGT-TNNI3K*	G	A	0.024	0.99438	Yes
rs17024393	1	*GNAT2*	C	T	0.066	0.98606	Yes
rs2820292	1	*NAV1*	C	A	0.020	1	Yes
rs3101336	1	*NEGR1*	C	T	0.033	1	Yes
rs543874	1	*SEC16B*	G	A	0.048	0.98967	Yes
rs657452	1	*AGBL4*	A	G	0.023	0.98647	Yes
rs1016287	2	*LINC01122*	T	C	0.023	0.993	Yes
rs10182181	2	*ADCY3*	G	A	0.031	0.99444	Yes
rs11126666	2	*KCNK3*	A	G	0.021	0.99398	Yes
rs11688816	2	*EHBP1*	G	A	0.017	0.9801	Yes
rs13021737	2	*TMEM18*	G	A	0.060	0.99085	Yes
rs1528435	2	*UBE2E3*	T	C	0.018	0.99703	Yes
rs2121279	2	*LRP1B*	T	C	0.025	0.98629	Yes
rs7599312	2	*ERBB4*	G	A	0.022	0.97038	Yes
rs13078960	3	*CADM2*	G	T	0.030	0.9912	Yes
rs1516725	3	*ETV5*	C	T	0.045	0.99152	Yes
rs16851483	3	*RASA2*	T	G	0.048	0.99867	Yes
rs2365389	3	*FHIT*	C	T	0.020	0.99235	Yes
rs3849570	3	*GBE1*	A	C	0.019	0.99521	Yes
rs6804842	3	*RARB*	G	A	0.019	0.98662	Yes
rs10938397	4	*GNPDA2*	G	A	0.040	1	Yes
rs11727676	4	*HHIP*	T	C	0.036	1	Yes
rs13107325	4	*SLC39A8*	T	C	0.048	1	Yes

Table continued

rs17001654	4	*SCARB2*	G	C	0.031	N/A	No, not in HWE and missing genotype data more than 5%
rs2112347	5	*POC5*	T	G	0.026	1	Yes
rs13191362	6	*PARK2*	A	G	0.028	0.98899	Yes
rs2033529	6	*TDRG1*	C	A	0.019	N/A	No, SNP not available
rs205262	6	*C6orf106*	G	A	0.022	0.99654	Yes
rs2207139	6	*TFAP2B*	G	A	0.045	0.99881	Yes
rs9400239	6	*FOXO3*	C	T	0.019	0.99206	Yes
rs1167827	7	*HIP1*	G	A	0.020	1	Yes
rs2245368	7	*PMS2L11*	C	T	0.032	1	Yes
rs17405819	8	*HNF4G*	T	C	0.022	0.99776	Yes
rs2033732	8	*RALYL*	C	T	0.019	1	Yes
rs10733682	9	*LMX1B*	C	G	0.017	N/A	No, missing genotype data more than 5%
rs10968576	9	*LINGO2*	G	A	0.025	1	Yes
rs1928295	9	*TLR4*	T	C	0.019	0.99997	Yes
rs4740619	9	*C9orf93*	T	C	0.018	0.99705	Yes
rs6477694	9	*EPB41L4B*	C	T	0.017	0.98899	Yes
rs11191560	10	*NT5C2*	C	T	0.031	0.99976	Yes
rs17094222	10	*HIF1AN*	C	T	0.025	0.96755	Yes
rs7899106	10	*GRID1*	C	A	0.040	0.98298	Yes
rs11030104	11	*BDNF*	A	G	0.041	0.99924	Yes
rs12286929	11	*CADM1*	G	A	0.022	0.98996	Yes
rs2176598	11	*HSD17B12*	T	C	0.020	1	Yes
rs3817334	11	*MTCH2*	T	C	0.026	1	Yes
rs4256980	11	*TRIM66*	G	C	0.021	0.99239	Yes
rs11057405	12	*CLIP1*	C	A	0.031	1	Yes
rs7138803	12	*BCDIN3D*	A	G	0.032	1	Yes
rs12429545	13	*OLFM4*	A	G	0.033	0.97383	Yes

Table continued

rs9581854	13	*MTIF3*	T	C	0.030	0.98642	Yes (SNP reported in GIANT analysis (rs12016871) has been renamed rs9581854)
rs10132280	14	*STXBP6*	C	A	0.023	0.97431	Yes
rs11847697	14	*PRKD1*	T	C	0.049	1	Yes
rs12885454	14	*PRKD1*	C	A	0.021	0.99533	Yes
rs7141420	14	*NRXN3*	T	C	0.024	0.97984	Yes
rs16951275	15	*MAP2K5*	T	C	0.031	0.99791	Yes
rs3736485	15	*DMXL2*	A	G	0.018	0.98642	Yes
rs12446632	16	*GPRC5B*	G	A	0.040	0.99968	Yes
rs1558902	16	*FTO*	A	T	0.082	0.99912	Yes
rs2650492	16	*SBK1*	A	G	0.021	0.98003	Yes
rs3888190	16	*ATP2A1*	A	C	0.031	0.99791	Yes
rs758747	16	*NLRC3*	T	C	0.023	N/A	No, missing genotype data more than 5%
rs9925964	16	*KAT8*	C	G	0.019	N/A	No, not in HWE
rs1000940	17	*RABEP1*	G	A	0.019	0.9958	Yes
rs12940622	17	*RPTOR*	G	A	0.018	0.99774	Yes
rs1808579	18	*C18orf8*	C	T	0.017	0.99769	Yes
rs6567160	18	*MC4R*	C	T	0.056	0.9957	Yes
rs7243357	18	*GRP*	T	G	0.022	0.98911	Yes
rs17724992	19	*PGPEP1*	A	G	0.019	0.98244	Yes
rs2075650	19	*TOMM40*	C	G	0.026	N/A	No, not in HWE
rs2287019	19	*QPCTL*	C	T	0.036	0.97714	Yes
rs29941	19	*KCTD15*	G	A	0.018	1	Yes
rs3810291	19	*ZC3H4*	A	G	0.028	1	Yes

Abbreviations: Chromosome (Chr.), Hardy-Weinberg equilibrium (HWE)

Table 7.2: Characteristics of study participants

Characteristic	Number of women
Number of participants	46,784
Age (years), mean (SD)	56.1 (7.9)
BMI at baseline (kg/m^2), median (IQR)	26.0 (23.4-29.5)
Genetic risk score, mean (SD)	31.4 (2.6)
Physical activity (MET-hours/week), median (IQR)	29.3 (13.5-58.2)
Leisure screen time (hours/day), mean (SD)	4.1 (6.5)

Abbreviations: body mass index (BMI), interquartile range (IQR), metabolic equivalent (MET), standard deviation (SD)

Table 7.3: Association of fifths of genetic risk score with body mass index (kg/m^2)

GRS fifths	N	GRS range	GRS mean (SD)	Mean accelerometer value (SD) (milli-gravity)	Mean BMI (95% CI) (kg/m^2)
1	9,489	20.8 -	27.8 (1.2)	28.6 (8.1)	25.8 (25.7-25.9)
2	9,402	29.2 -	30.0 (0.4)	28.4 (7.9)	26.2 (26.1-26.2)
3	9,369	30.7 -	31.4 (0.4)	28.4 (8.4)	26.5 (26.5-26.6)
4	9,312	32.1 -	32.8 (0.4)	28.5 (8.1)	26.8 (26.7-26.9)
5	9,212	33.6-41.5	35.1 (1.2)	28.0 (7.6)	27.3 (27.2-27.4)
					P for trend < 0.0001

Abbreviations: body mass index (BMI), confidence interval (CI), genetic risk score (GRS), standard deviation (SD)
Analysis was adjusted for age, age^2, region of recruitment, and self-reported health status

Table 7.4: Adjusted geometric means and (95% confidence intervals) of body mass index (kg/m^2) by fifths of genetic risk score and physical activity levels

	Body mass index (kg/m^2)		
Fifths of genetic risk score	Low physical activity (<10.0 MET-hours/week)	Moderate physical activity (<10-49.9 MET-hours/week)	High physical activity (≥50.0 MET-hours/week)
1	26.5 (26.3-26.7)	25.8 (25.7-25.9)	25.5 (25.3-25.6)
2	27.2 (27.0-27.4)	26.1 (26.0-26.2)	25.6 (25.5-25.8)
3	27.5 (27.3-27.7)	26.5 (26.4-26.6)	26.1 (25.9-26.2)
4	28.0 (27.8-28.2)	26.7 (26.6-26.8)	26.3 (26.1-26.5)
5	28.6 (28.3-28.8)	27.2 (27.1-27.3)	26.7 (26.5-26.8)

Table 7.5: Association of fifths of genetic risk score with body mass index (kg/m^2), by physical activity level

Physical activity level [a]	N	MET-hours/week	Mean accelerometer PA[b] (SD) (milli-gravity)	β[c] (SE)	β[d] (SE)	$N_{normal\ weight}/N_{obese}$	OR[e] (SE)	P
All women	46,784		28.4 (8.0)	0.218 (0.008)	1.60 (0.070)	24,928/15,263	2.22 (2.05-2.41)	<0.0001
Low	8,727	<10.0	25.9 (7.0)	0.299 (0.023)	2.23 (0.192)	2,730/2,885	2.50 (2.09-2.99)	<0.0001
Moderate	23,900	10-49.9	28.1 (7.8)	0.207 (0.012)	1.53 (0.096)	9,929/5,128	2.20 (1.96-2.47)	<0.0001
High	14,157	≥50.0	30.4 (8.5)	0.188 (0.014)	1.33 (0.118)	6,564/2,531	2.20 (1.88-2.56)	<0.0001
				P for interaction = 0.001				

Abbreviations: body mass index (BMI), metabolic equivalent (MET), odds ratio (OR), physical activity (PA), standard deviation (SD), standard error (SE)

[a] Total physical activity levels were categorised according to the International Physical Activity Questionnaire (IPAQ) guidelines

[b] Mean objective physical activity measures are based on overall average acceleration during a one week period, measured in units of milli-gravity

[c] Increase in BMI (kg/m^2) for each additional unit increase in the genetic risk score

[d] Increase in BMI (kg/m^2) comparing top to reference bottom category of genetic risk score

[e] Increase in the odds of being obese versus being normal weight, comparing the top to reference bottom category of genetic risk score

Table 7.6: Association of fifths of physical activity with body mass index (kg/m^2), by genetic risk score

				Change in BMI by self-reported physical activity	
Category of genetic risk [a]	N	GRS Range	Mean GRS (SD)	B[b] (SE)	B[c] (SE)
All women	10,971	20.8-40.6	31.4 (2.6)	-0.395 (0.031)	-1.67 (0.140)
Low genetic risk score	3,661	20.8-30.2	28.5 (1.3)	-0.345 (0.050)	-1.46 (0.221)
Medium genetic risk score	3,654	30.2-32.5	31.4 (0.65)	-0.393 (0.053)	-1.59 (0.238)
High genetic risk score	3,656	32.5-40.6	34.2 (1.4)	-0.445 (0.059)	-1.89 (0.261)
				P for interaction = 0.672	
				Change in BMI by accelerometer-measured physical activity	
Category of genetic risk [a]	N	GRS Range	Mean GRS (SD)	B[b] (SE)	B[c] (SE)
All women	10,971	20.8-40.6	31.4 (2.6)	-0.778 (0.032)	-3.28 (0.141)
Low genetic risk score	3,661	20.8-30.2	28.5 (1.3)	-0.644 (0.051)	-2.77 (0.224)
Medium genetic risk score	3,654	30.2-32.5	31.4 (0.65)	-0.730 (0.053)	-3.10 (0.235)
High genetic risk score	3,656	32.5-40.6	34.2 (1.4)	-0.975 (0.061)	-4.05 (0.270)
				P for interaction = 0.0069	

Abbreviations: body mass index (BMI), genetic risk score (GRS), standard deviation (SD), standard error (SE)

[a] Categories of genetic risk score were split by tertiles

[b] Change in BMI units (SD) (kg/m^2) for each additional fifth increase in physical activity (self-reported or accelerometer-measured)

[c] Change in BMI (kg/m^2) comparing top to reference bottom fifth of physical activity (self-reported or accelerometer-measured)

Table 7.7: Association of fifths of genetic risk score with body mass index (kg/m^2), by leisure screen time

Screen time per day [a]	N	β^b (SE)	β^c (SE)	β^d (SE)	$n_{normal\ weight}/n_{obese}$	OR^e (SE)	P
All women	46,784	0.213 (0.008)	0.214 (0.008)	1.56 (0.069)	19,223/10,544	2.24 (2.06-2.43)	<0.0001
≤ 2 hours	11,298	0.190 (0.015)	0.191 (0.015)	1.39 (0.127)	6,155/1,571	2.40 (1.99-2.90)	<0.0001
2.5 - 4 hours	21,752	0.194 (0.012)	0.196 (0.012)	1.45 (0.010)	8,981/4,611	2.11 (1.87-2.37)	<0.0001
> 4 hours	13,734	0.262 (0.017)	0.263 (0.017)	1.88 (0.142)	4,087/4,362	2.35 (2.04-2.72)	<0.0001
			P for interaction = 0.016				

Abbreviations: body mass index (BMI), odds ratio (OR), standard error (SE)

[a] Screen time consists of television watching and leisure time computer use

[b] Increase in BMI units (kg/m^2) for each additional unit increase in the genetic risk score

[c] further adjusted for physical activity

[d] Increase in BMI (kg/m^2) comparing top to reference bottom category of genetic risk score

[e] Increase in the odds of being obese versus being normal weight comparing the top to reference bottom category of genetic risk score

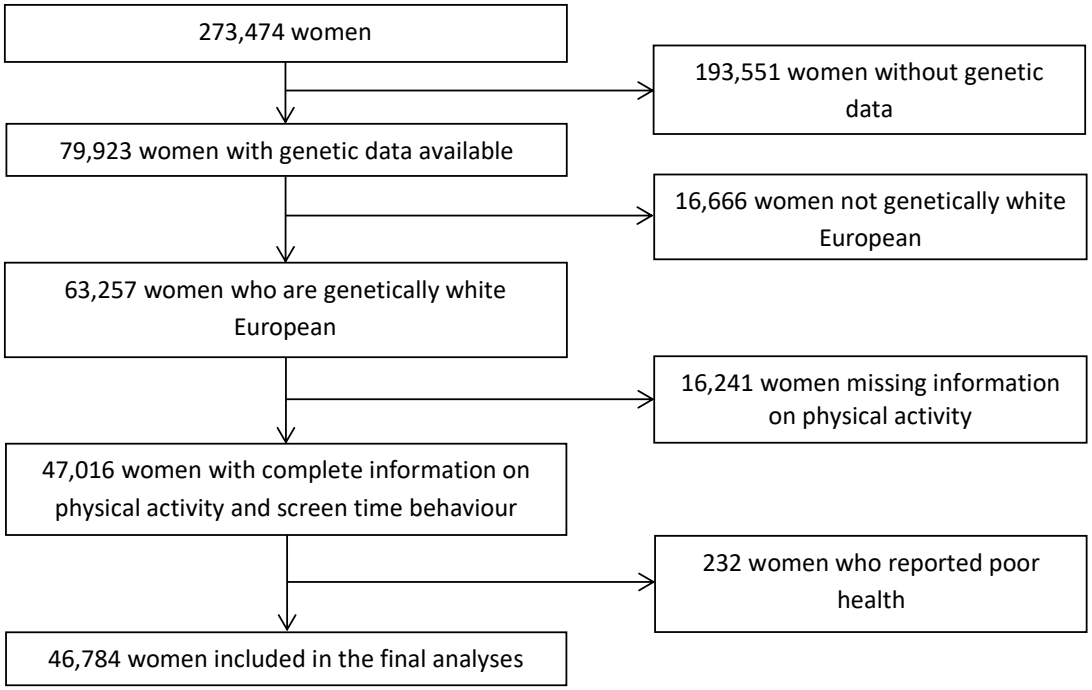
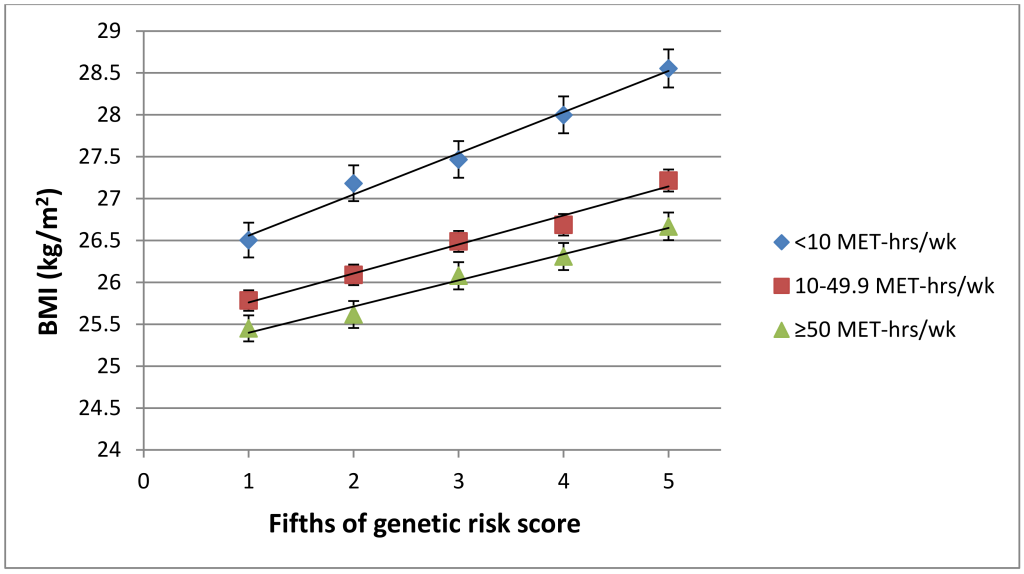
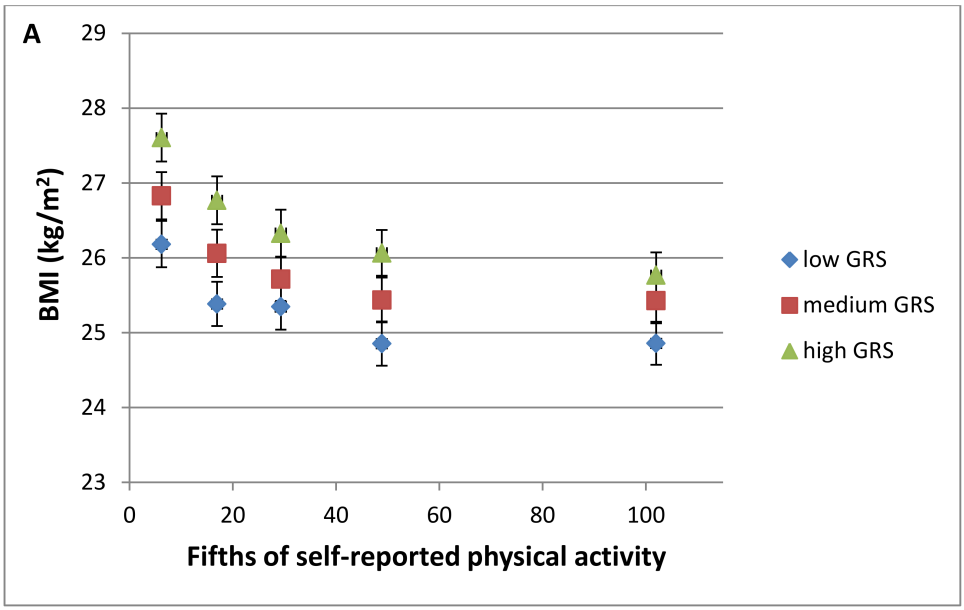


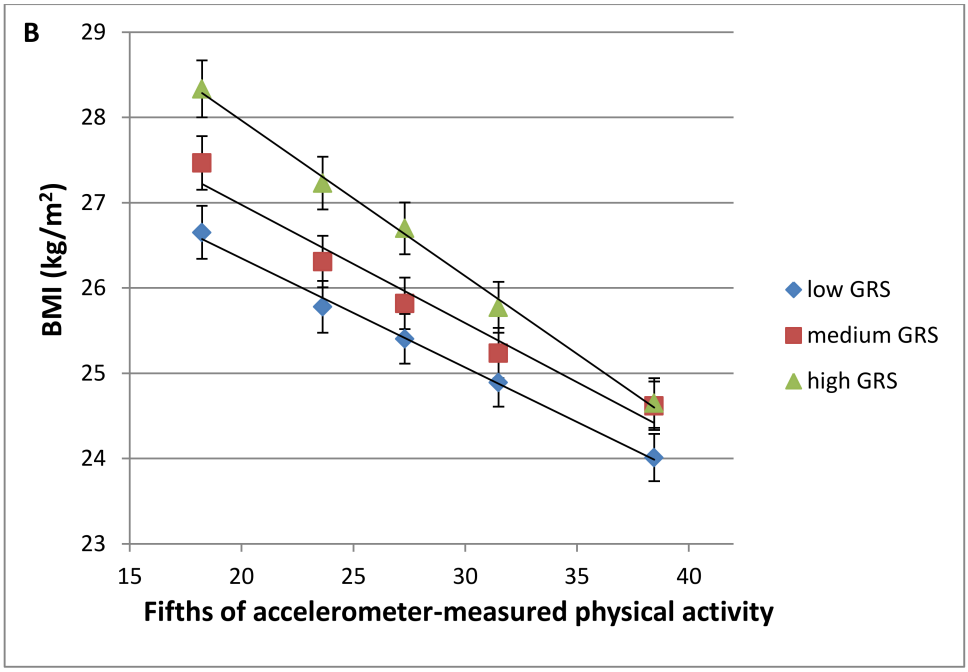
Figure 7.1: Participant inclusion/exclusion flow diagram for analyses of gene-environment interactions in adiposity



Point estimates and 95% confidence intervals are shown.

Figure 7.2: Association of fifths of genetic risk score (based on quintile cut-points) and body mass index (BMI), by physical activity level in metabolic equivalent hours per week (MET-hrs/wk)

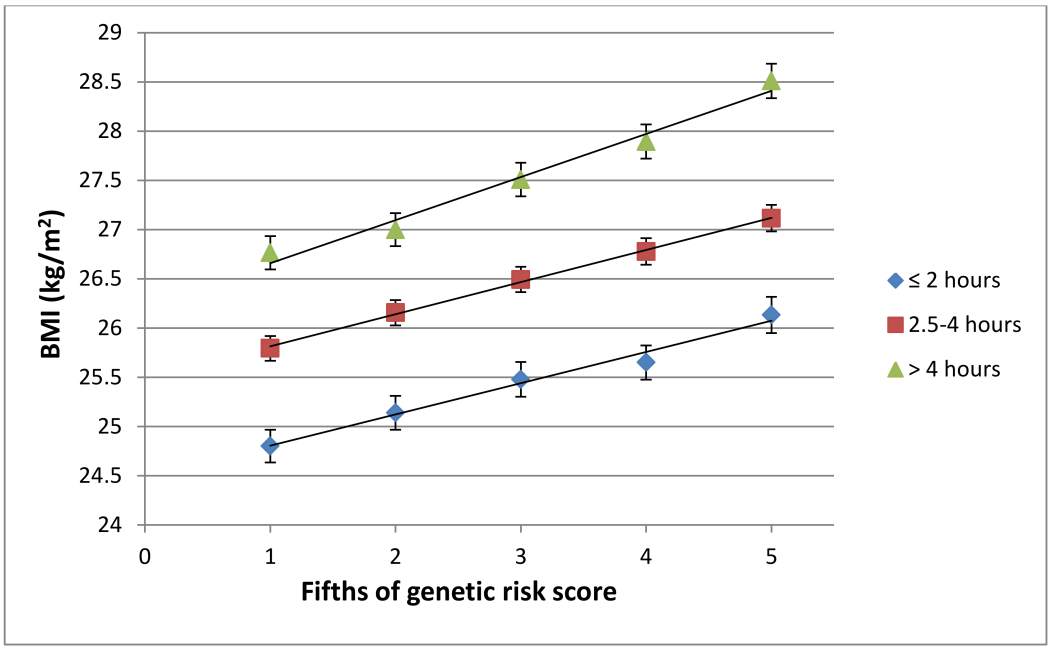




Low genetic risk score represents the group with the lowest genetic predisposition to adiposity and high genetic risk score represents the group with the highest genetic predisposition to adiposity, as captured by the weighted genetic risk score.

Point estimates and 95% confidence intervals are shown.

Figure 7.3: Association of physical activity, measured by A) self-reported questionnaire and B) accelerometer, with body mass index (BMI) by thirds of genetic risk score (GRS)



Screen time behaviour is the sum of television watching and leisure computer use time.

Point estimates and 95% confidence intervals are shown.

Figure 7.4: Association of fifths of genetic risk score and body mass index (BMI), by screen time behaviour

Chapter 8

INDICES OF ADIPOSITY AND BREAST CANCER RISK IN POSTMENOPAUSAL WOMEN

8.1 Introduction

Although there are several important known risk factors for breast cancer, excess adiposity is one of the few modifiable ones. Previous studies have demonstrated an association between obesity and increased risk for postmenopausal breast cancer,[142] but questions remain about this relationship and its underlying mechanisms. Since adipose tissue is the major site of oestrogen synthesis in postmenopausal women, the most prominent hypothesis is that more body fat is linked to increased postmenopausal breast cancer risk through oestrogen-stimulated carcinogenesis[11,201]. For this reason, fat mass may be a better predictor of postmenopausal breast cancer risk than other measures of adiposity.

Due to high cost and difficulty of implementation, most large prospective cohort studies have been unable to directly address this question, instead examining the association between obesity and breast cancer risk using only traditional anthropometric measures such as body mass index (BMI), waist circumference, and waist-to-hip ratio, which do not directly distinguish between lean and fat mass[4,24,30,31]. While BMI is unable to capture changes in adiposity with age, bioelectrical impedance measures reflect increases in fat mass with age[5]. Furthermore, it is unclear whether the distribution of adipose tissue is relevant for breast cancer risk[6]. Most studies rely on self-reported measures of body size, which tend to underestimate adiposity[202]. The aim of this chapter is to clarify the association between obesity and postmenopausal breast cancer risk using data on bioelectrical impedance measures of body fat in UK Biobank.

8.2 Methods

8.2.1 Data source

Data were obtained from UK Biobank (reference number 3248, approved August 2013). Details of the rationale, design and survey methods for UK Biobank have been described elsewhere[110] and information on data available and access procedures are given on the study website (`http://www.ukbiobank.ac.uk/`). UK Biobank has approval from the North West Multi-centre Research Ethics Committee, the Confidentiality Advisory Group in England and Wales, and the Community Health Index Advisory Group in Scotland. All participants provided written informed consent.

8.2.2 Study participants

The complete UK Biobank dataset at the time of analyses for this chapter included 502,620 UK adults (229,165 men and 273,455 women) aged 40 to 69 at recruitment during 2006-2010.

Women were eligible for these analyses if they were postmenopausal at recruitment. Women were defined as being postmenopausal at recruitment if they reported that their periods had stopped. Women with unknown self-reported menopausal status were defined as postmenopausal if they were aged 53 or over at recruitment based on previously established criteria[84]. Women were excluded if they had a prior cancer diagnosis (except for non-melanoma skin cancer ICD-10 C44) (n=18,372), were premenopausal (n=62,899) or had unknown menopausal status after applying the above categorisation criteria (n=12,563), were current users of hormone replacement therapy (HRT) (n= 13,680), or had missing data on any measure of body size and composition (n=3,250). Current HRT users were excluded because HRT use is known to attenuate associations of adiposity with breast cancer risk in postmenopausal women[24,82,203]. A total of 162,691 postmenopausal women were included in the analyses (Figure 8.1).

8.2.3 Anthropometry and body composition

At the UK Biobank baseline interview, trained staff measured standing height using the Seca 202 device (Seca, Hamburg, Germany). The Wessex non-stretchable sprung tape measure (Wessex, United Kingdom) was used to measure waist circumference and hip circumference, from which waist-to-hip-ratio was derived by dividing waist circumference by hip circumference. The Tanita BC-418MA body

composition analyser (Tanita, Tokyo, Japan) was used to measure body and trunk fat mass and percentage using bio-impedance. BMI was calculated by dividing weight (kg) by the square of standing height (m^2). Dual-energy X-ray absorptiometry (DXA) was used to measure fat mass/percentage on a subset of participants beginning in 2014 using the GE-Lunar iDXA (GE Healthcare, Chicago, USA).

8.2.4 Ascertainment of cancer cases

UK Biobank obtains data on cancer diagnoses through NHS Digital for participants in England and Wales, and NHS Central Register for participants in Scotland. It should be noted that UK Biobank does not cover Northern Ireland, despite its UK name.

Additionally, although I was not personally involved in data collection for UK Biobank, it is important to consider the quality of the data and possible limitations in the flagging of cancer registry data. This is an important consideration since inaccurate and incomplete data can seriously bias estimates of cancer incidence and survival. Errors can arise due to coding discrepancies or discordant diagnoses, for example, if one diagnosis is made on the hospital discharge file but a different diagnosis is made after histological analysis[204].

Completeness of case ascertainment in English cancer registries is reported to be around 98-99%, based on a study that linked routine cancer registration with information from the Hospital Episode Statistics database for the years 2001-2007[127]. Despite my best efforts, it is difficult to find concrete information on the process of flagging cancer cases for UK Biobank. There is a variable in UK Biobank that provides reason for loss to follow-up. This variable was not available

in my analysis dataset. However, counts of participants are available on the UK Biobank website and are informative for better understanding the limitations of follow-up data. A total of 1,299 participants were lost to follow up, as of 23 April 2018. The top three reasons for loss to follow up were NHS records indicating that 604 participants had left the UK, UK Biobank sources indicating that 496 participants had left the UK, and 158 participants having withdrawn consent for future linkage.

8.2.5 Statistical analyses

Baseline characteristics of study participants were compared between top and bottom quartiles of BMI, body fat mass, and waist circumference. Associations between the various measures of body size and composition were examined by calculating age-adjusted Pearson's partial correlation coefficients.

Women were followed from date of baseline assessment centre visit until the earliest of: date of breast cancer registration (ICD-10 C50), date of death, date of loss to follow-up or end of follow-up for cancer incidence (30 November 2014). Women diagnosed with any cancer other than breast cancer (with the exception of non-melanoma skin cancer) during follow-up were censored at date of diagnosis.

Multivariable-adjusted Cox regression with attained age as the underlying time variable was used to estimate hazards ratios (referred to as relative risks [RRs]) and 95% confidence intervals (CIs) for the association between breast cancer risk and body size and composition measures: height, weight, BMI, body fat mass, body fat percentage, waist circumference, hip circumference, waist-to-hip ratio, trunk fat mass, and trunk fat percentage. It should be noted that floating absolute risks

were used to present the relative risk estimates in Chapter 9 with the rationale of avoiding problems associated with choosing an arbitrary reference group. However, for the analyses presented in the current chapter, floating absolute risks were not used, because I felt that the lowest category of adiposity was a natural reference group. For example, for BMI, the lowest quartile with a median BMI in the healthy range was considered a natural reference group for the analyses. For the risk analyses, women were categorised into quartiles according to their baseline body size/composition. To facilitate comparison across measures of adiposity, I also estimated RRs per standard deviation (SD) increase and show the results from tests for linear trend with categories recoded using the repeat assessment median within each category (grouped into quartiles) and the corresponding χ^2 statistic. Tests for linear trend were performed with categories coded in an ordinal fashion using the repeat assessment median of each category in order to attain measures that are more representative of the true long-term values[205]. In a separate analysis, tests for linear trend were performed using the DXA median of each quartile for fat mass/percentage variables. All analyses were stratified by 5-year age at recruitment categories, region of recruitment, and socioeconomic status (based on fifths of Townsend deprivation index)[119]. All analyses were adjusted for family history of breast cancer (no, yes), age at menarche ($<$12, 12-13, $\geq$14 years), age at first birth ($<$25, 25-29, $\geq$30 years), parity (nulliparous, 1-2, $\geq$3), age at menopause ($<$45, 45-54, $\geq$55 years), previous HRT use (never, past) smoking (never, past, current), alcohol intake frequency (less than 3 times a month, 1-4 times a week, daily or almost daily), physical activity ($<$14, 14-30, 31-59.8, $\geq$59.9 metabolic equivalent hours per week), and ethnicity (White, Mixed, Asian or Asian British, Black or Black British, Chinese, other ethnic group). Regression models

with exposures other than height as the exposure of interest were adjusted for height (<160, 160-165.9, ≥166 cm) as an independent risk factor for breast cancer. Analyses were further adjusted for body size at age 10 (thinner, about average, plumper) in order to account for the independent association between greater childhood body fatness and decreased breast cancer risk[34]. Women with missing values for any of the adjustment variables were assigned to a separate "unknown" category for the respective variable. Information was either missing or reported as unknown for <3% of covariates, with the exception of total physical activity (28.5%) and age at menopause (18.7%) which may be difficult to report accurately if menopausal status is masked by hysterectomy or HRT before menstrual periods stop naturally[84]. To assess the impact of missing values, I conducted a sensitivity analysis restricted to participants with known values for all adjustment variables.

The χ^2 statistic for trend across quartile medians was used to quantify the extent to which each particular index of adiposity was related to risk[206]. I further adjusted anthropometric measures of central adiposity by the impedance measure of body fat mass, but chose not to mutually adjust impedance measures for each other, because these variables were highly correlated (r >0.88). Assessment of interaction terms between each exposure of interest and the underlying time variable did not suggest any significant deviation from proportional hazards.

RRs were also estimated for each body size/composition variable as a function of years since menopause (<12, ≥12 years) by using a time-varying covariate. Comparisons were made between women who were <12 years and ≥12 years since menopause because this cut point divided the number of cases approximately equally. Likelihood ratio tests were used to assess whether time since menopause was an effect modifier in the associations between body size/composition variables

and breast cancer risk. All analyses were conducted using STATA version 15.0 (Stata Corp LP, College Station, TX).

8.3 Results

Two thousand nine hundred and thirteen invasive breast cancer cases were diagnosed among 162,691 postmenopausal women at risk during a mean follow-up of 5.7 (SD 1.1) years. Table 8.1 shows baseline characteristics of participants according to BMI, body fat mass, and waist circumference. Women in the top quartiles for these body size/composition variables were more likely to be in the lowest fifth of socioeconomic class, have a younger age at first birth, have used HRT, consume alcohol only occasionally or never, and engage in less physical activity while the leanest women were more likely to have used oral contraceptives in the past. Those in the bottom quartiles of BMI and body fat mass were more likely to be either never or current smokers.

Table 8.2 shows age-adjusted Pearson's partial correlation coefficients for the various anthropometric and body composition measures. BMI was highly correlated with both body fat percentage (0.85) and body fat mass (0.94). Total and central adiposity were very highly correlated as shown by the correlation coefficients between body and trunk fat percentage (0.98) and body and trunk fat mass (0.97). In view of these high correlations between certain indices, associations of risk with overall fatness were not mutually adjusted for other indices of body fatness. Results for both body and trunk fat mass are presented, but the extremely high correlation coefficients observed between these two indices indicate that their effects may be indistinguishable.

All measures of overall adiposity, as well as height, were associated with increased postmenopausal breast cancer risk in an approximately linear fashion (Figure 8.2, Table 8.3). After adjusting for body size at age 10, the magnitudes of association for all indices of adiposity were increased. Lowest breast cancer risk was associated with being in the lowest quartile of adiposity (median BMI 22.2 kg/m^2). Women in the top quartile of body fat mass (32.6-108.4 kg; repeat assessment median, 37.6 kg), had the greatest increase in breast cancer risk (RR, 1.70; 95% CI, 1.52-1.90; p-trend $<$0.001), compared with women in the lowest quartile (8.0-20.5 kg; repeat assessment median, 17.6 kg). The attributable risk for body fat mass for breast cancer in the top quartile expressed as a fraction was 0.38. This was calculated as (incidence in the exposed - incidence in the unexposed)/incidence in the exposed. The magnitudes of associations with breast cancer risk per SD increases in BMI and body fat mass were similar with overlapping confidence intervals: RR 1.21; 95% CI 1.15-1.27 for BMI and RR 1.25; 95% CI 1.19-1.31 for body fat mass in multivariable-adjusted models (Tables 8.3 and 8.4). Recoding quartiles using the corresponding DXA instead of repeat assessment medians for tests of linear trend produced nearly identical results (Table 8.5).

All indicators of central obesity were significantly associated with postmenopausal breast cancer risk (Figure 8.2, Table 8.4); fat mass was slightly more informative about postmenopausal breast cancer risk (χ^2=259.4), and waist-to-hip ratio was least informative (χ^2=202.0). After adjusting for body fat mass, associations of breast cancer risk with per SD increases in waist circumference, hip circumference, and waist-to-hip ratio were attenuated to the null. However, the confidence intervals around the adjusted associations were moderately wide so small residual associations cannot be ruled out. The results of sensitivity analyses excluding par-

ticipants with any missing values did not differ materially from the main findings.

The magnitudes of risk associated with all measures of adiposity were greater in women who had been postmenopausal for 12 or more years; p values for interaction were <0.02 for all adiposity variables except waist-to-hip ratio, BMI, and body and trunk fat percentage (Table 8.6). For instance, comparing the top to the bottom quartile of body fat mass, postmenopausal women who were at least 12 years since menopause had a greater increased risk of breast cancer (RR, 1.86; 95% CI, 1.56-2.21) compared to those who were less than 12 years since menopause (RR, 1.47; 95% CI 1.24-1.75).

8.4 Discussion

In this large prospective study of postmenopausal women, there were statistically significant positive associations between all measures of adiposity and breast cancer risk. The relative risks, comparing the top with the bottom quartiles, were marginally greater for fat mass measured by impedance, but the RRs per SD increase and χ^2 statistics associated with each index of adiposity were of similar magnitude. All measures of central adiposity were also associated with postmenopausal breast cancer risk but these associations were attenuated after adjusting for body fat mass. Associations across all measures of adiposity were stronger in women who had been postmenopausal for $\geq$12 years.

Previous studies using impedance measures (based on 12,159; 7,523; and 13,598 participants) demonstrated a moderately greater magnitude of risk associated with postmenopausal breast cancer when comparing body fat to its surrogate measure, BMI[207–209]. Another study using DXA measures with 503 incident breast cancer

cases concluded that there is no difference in the ability to predict postmenopausal breast cancer risk when comparing anthropometric indices to DXA-derived measures of body fat[210]. The results presented here are consistent with the findings from the latter study in that BMI and waist circumference were as informative for breast cancer risk as impedance measures of adiposity.

In these data, all measures of adiposity showed stronger associations with postmenopausal breast cancer risk among women who were $\geq$12 years since menopause compared to women with $<$12 years since menopause. These findings are consistent with those from two previous large prospective studies which found that the positive association between adiposity and postmenopausal breast cancer risk was more marked in women aged 65 or older[31] and in women who had experienced menopause $\geq$15 years previously[209].

Elevated oestrogen levels may stimulate carcinogenesis through increased cell proliferation as well as through pro-angiogenic and anti-apoptotic effects[201,211]. The relationship of adiposity with breast cancer risk in postmenopausal women can largely be explained by increased endogenous oestrogen and decreased sex hormone-binding globulin. In two large studies with prospectively measured oestrogen, adjustment for free estradiol completely attenuated the association between adiposity and breast cancer risk, suggesting that this relationship may be mediated by the concentration of bioavailable estradiol[22,212]. The results presented in this chapter further suggest that the excess risk associated with increased levels of endogenous oestrogens may take several years to fully outweigh the reduction in risk associated with adiposity in premenopausal women[203,213].

A 2003 review on central obesity found that waist circumference and waist-to-hip ratio were no longer associated with postmenopausal breast cancer risk after

adjusting for BMI or weight,[214] and some subsequent studies showed similar results[203,215,216]. However, a recent meta-analysis found that the association between waist circumference and postmenopausal breast cancer risk was only slightly attenuated after adjusting for BMI[6]. In the present analyses, which were adjusted for measured body fat mass rather than BMI, all associations between postmenopausal breast cancer and anthropometric measures of central adiposity were attenuated after adjustment for overall body fat, suggesting that overall body fat is the underlying adiposity-related predictor of postmenopausal breast cancer risk.

Recent Mendelian randomization studies report an inverse association between BMI predicted using single nucleotide polymorphisms identified by genome-wide association studies and postmenopausal breast cancer risk[38,39]. However, the BMI genetic score used in these studies may be more predictive of BMI in early-life rather than later adulthood[40]. In agreement with prior studies,[34,217] the results in this thesis, as shown in Chapter 4, demonstrate a protective association between greater childhood adiposity and breast cancer risk. I subsequently adjusted for comparative body size at age 10 as an independent risk factor for breast cancer and present relative risks both before and after this adjustment.

A notable strength of the present study is the availability of objectively-measured, rather than self-reported, adult body size and composition measures. Self-reported measures of adiposity such as weight tend to be underestimated, especially among overweight and obese women[202]. Other strengths include the prospective design, large sample size, and virtually complete follow-up[127]. The availability of a wide range of known and putative risk factors for breast cancer allowed potential bias and confounding to be addressed.

Bioelectrical impedance measurements have been validated against hydroden-

sitometry, which is regarded as the gold standard for measuring body fat[218,219]. Although bioelectrical impedance measurements capture body fat more accurately than anthropometric measurements,[5,220] they have limitations. Studies assessing bioelectrical impedance in diverse populations may require population-specific calibration equations for different ethnic groups[5]. This is of limited concern in the present study since over 95% of participants are White and ethnic group is adjusted for in the analyses in this chapter. Hydration status can also affect impedance analysis[221]. However, the accuracy of body composition measurements was improved by using a standardised measurement protocol administered by trained staff as well as utilising repeat assessment data on all the body size and composition measures. Another limitation of the study is the lack of information on hormone receptor status of the tumor (especially information on oestrogen receptor); however since 70-80% of breast cancers included in this present study of postmenopausal women are likely to be oestrogen receptor positive, the results are largely representative of oestrogen receptor positive breast cancers tumours[222].

8.5 Conclusion

In this large prospective study, I found strong positive associations between all body size/composition variables and postmenopausal breast cancer risk. Anthropometric indices, with the exception of waist-to-hip ratio, were as informative for breast cancer risk as impedance measures of fat mass. The results also suggest no association between postmenopausal breast cancer risk and abdominal adiposity beyond its contribution to overall adiposity.

Table 8.1: Baseline characteristics for the 162,691 postmenopausal participants of UK Biobank by various measures of adiposity

	Body mass index (kg/m^2)		Body fat mass (kg)		Waist circumference (cm)	
	Quartile 1	**Quartile 4**	**Quartile 1**	**Quartile 4**	**Quartile 1**	**Quartile 4**
Age at recruitment, mean (SD)	59.8 (5.4)	60.3 (5.4)	59.9 (5.4)	60.3 (5.4)	59.6 (5.4)	60.6 (5.3)
Lowest socioeconomic quintile, %	16	25.1	16.7	23.8	15.2	25.5
Family history of breast cancer, %	7.2	6.1	6.8	6.3	7	6.2
Age at menarche, mean (SD)	13.2 (1.6)	12.7 (1.7)	13.1 (1.6)	12.7 (1.7)	13.1 (1.6)	12.7 (1.7)
Age at first birth, mean (SD)	26.4 (4.9)	24.6 (4.8)	26.2 (4.9)	24.7 (4.9)	26.2 (4.9)	24.7 (4.9)
Parity, mean (SD)	1.8 (1.2)	2.0 (1.3)	1.8 (1.2)	2.0 (1.2)	1.8 (1.1)	2.0 (1.3)
Ever oral contraceptive use, %	78.8	75.3	78.3	76.2	79	75.1
Age at menopause, mean (SD)	49.9 (4.6)	49.6 (5.5)	49.9 (4.7)	49.6 (5.4)	49.9 (4.6)	49.6 (5.4)
Past hormone replacement therapy use, %	43.1	47.6	42.6	47.8	42.7	47.5
Smoking, %						
Never	61	56.5	61.1	55.7	62.6	54.5
Past	29.7	35.6	29.3	36.5	29.1	36.6
Current	9.1	7.4	9.4	7.2	8	8.4
Alcohol intake frequency, %						
Special occasions only or never	21.1	36.2	22.2	34.4	21.3	35.8
1 to 3 times a month	10.3	15.1	10.6	14.8	11	14.4
1 to 4 times a week	46.3	38.4	45.8	39.6	47.5	38
Daily or almost daily	22.2	10.3	21.3	11.2	20.1	11.7
Height (cm), mean (SD)	162.9 (6.2)	160.8 (6.1)	160.5 (6.2)	162.9 (6.1)	161.5 (6.1)	162.0 (6.2)
Total physical activity (MET-hrs), mean (SD)	53.3 (56.8)	40.3 (50.2)	55.1 (58.7)	39.0 (48.4)	53.9 (57.0)	39.6 (49.3)

Table 8.2: Age-adjusted Pearson's partial correlation coefficients for the 162,691 postmenopausal participants

	Height, cm	Weight, kg	BMI, kg/m^2	Waist circumference, cm	Hip circumference, cm	Waist-to-hip ratio	Body fat mass, kg	Body fat %	Trunk fat mass, kg	Trunk fat %
Height, cm	1.00									
Weight, kg	0.27	1.00								
BMI, kg/m^2	-0.13	0.92	1.00							
Waist circumference, cm	0.04	0.86	0.87	1.00						
Hip circumference, cm	0.12	0.91	0.89	0.82	1.00					
Waist-to-hip ratio	-0.07	0.41	0.45	0.74	0.23	1.00				
Body fat mass, kg	0.15	0.97	0.94	0.88	0.91	0.43	1.00			
Body fat %	0.01	0.83	0.85	0.81	0.81	0.44	0.93	1.00		
Trunk fat mass, kg	0.24	0.92	0.85	0.83	0.86	0.41	0.97	0.94	1.00	
Trunk fat %	0.13	0.77	0.75	0.74	0.75	0.39	0.88	0.98	0.94	1.00

Table 8.3: Association of height and overall adiposity measures with invasive breast cancer risk among never and former postmenopausal hormone replacement therapy users

					Stratified only	Multivariable-adjusted	Further adjusted for body size at age 10	Further adjusted for body size at age 10
Height quartiles	**Range (cm)**	**Median (cm)**	**Repeat median (cm)**	**Cases *n***	**RR (95% CI)[a]**	**RR (95% CI)[b]**	**RR (95% CI)[d]**	**RR (95% CI)[d]**
1	121-158	155	155	800	1	1	1	**Per SD increase**
2	158.1-162	161	160	720	1.07 (0.97-1.18)	1.07 (0.96-1.18)	1.07 (0.97-1.18)	1.10 (1.05-1.15)
3	162.1-166	164	164	653	1.09 (0.98-1.20)	1.08 (0.97-1.20)	1.08 (0.97-1.20)	
4	166.1-193	169	169	740	1.30 (1.18-1.44)	1.28 (1.16-1.42)	1.27 (1.15-1.41)	
χ^2								175.7
Weight quartiles	**Range (kg)**	**Median (kg)**	**Repeat median (kg)**	**Cases *n***	**RR (95% CI)[a]**	**RR (95% CI)[c]**	**RR (95% CI)[d]**	**RR (95% CI)[d]**
1	34.0-61.9	57.4	57.1	565	1	1	1	**Per SD increase**
2	62.0-69.2	65.6	65.4	715	1.27 (1.14-1.42)	1.25 (1.12-1.40)	1.27 (1.13-1.42)	1.24 (1.18-1.30)
3	69.3-78.5	73.4	72.6	738	1.31 (1.18-1.47)	1.29 (1.15-1.44)	1.32 (1.18-1.48)	
4	78.6-179.0	86.9	85.6	895	1.61 (1.45-1.79)	1.58 (1.41-1.76)	1.66 (1.48-1.86)	
χ^2								254.9
BMI quartiles	**Range (kg/m^2)**	**Median (kg/m^2)**	**Repeat median (kg/m^2)**	**Cases *n***	**RR (95% CI)[a]**	**RR (95% CI)[c]**	**RR (95% CI)[d]**	**RR (95% CI)[d]**
1	14.6-23.7	22.2	22.2	579	1	1	1	**Per SD increase**
2	23.8-26.4	25.1	25.1	743	1.25 (1.12-1.39)	1.26 (1.13-1.41)	1.28 (1.15-1.43)	1.21 (1.15-1.27)
3	26.5-29.9	28	28	753	1.29 (1.15-1.43)	1.32 (1.18-1.47)	1.35 (1.21-1.51)	
4	30.0-68.1	33.1	33	838	1.45 (1.30-1.62)	1.52 (1.36-1.70)	1.60 (1.43-1.79)	
χ^2								242.6

Table continued

Body fat mass quartiles	Range (kg)	Median (kg)	Repeat median (kg)	Cases *n*	RR (95% CI)[a]	RR (95% CI)[c]	RR (95% CI)[d]	RR (95% CI)[d]
1	8.0-20.5	17.3	17.6	537	1	1	1	**Per SD increase**
2	20.6-25.8	23.2	23.4	727	1.33 (1.19-1.49)	1.31 (1.17-1.47)	1.32 (1.18-1.48)	1.25 (1.19-1.31)
3	25.9-32.5	28.8	28.8	756	1.37 (1.23-1.53)	1.35 (1.21-1.51)	1.38 (1.23-1.54)	
4	32.6-108.4	38.5	37.6	893	1.64 (1.48-1.83)	1.62 (1.45-1.82)	1.70 (1.52-1.90)	
χ^2								260.3
Body fat % quartiles	**Range (%)**	**Median (%)**	**Repeat median (%)**	**Cases *n***	**RR (95% CI)[a]**	**RR (95% CI)[c]**	**RR (95% CI)[d]**	**RR (95% CI)[d]**
1	9.4-32.9	29.7	30.5	550	1	1	1	**Per SD increase**
2	33.0-37.4	35.4	35.7	729	1.30 (1.16-1.45)	1.29 (1.15-1.44)	1.30 (1.16-1.45)	1.23 (1.17-1.29)
3	37.5-41.8	39.5	39.8	763	1.33 (1.19-1.48)	1.32 (1.18-1.48)	1.34 (1.20-1.50)	
4	41.9-69.8	44.8	44.6	871	1.54 (1.38-1.72)	1.55 (1.39-1.73)	1.60 (1.43-1.79)	
χ^2								249.4

[a] stratified by age at recruitment, region of recruitment, and socioeconomic status (Townsend deprivation index)

[b] adjusted for family history of breast cancer, age at menarche, age at first birth, parity, age at menopause, previous HRT use, smoking, alcohol intake frequency, physical activity, and ethnicity

[c] adjusted for family history of breast cancer, age at menarche, age at first birth, parity, age at menopause, previous HRT use, smoking, alcohol intake frequency, physical activity, ethnicity and height

[d] further adjusted for body size at age 10

Table 8.4: Association of central adiposity measures and invasive breast cancer risk among never and former post-menopausal HRT users

					Stratified only	Multivariable-adjusted	Further adjusted for body size at age 10	Further adjusted for body size at age 10	Further adjusted for body fat mass
Waist quartiles	**Range (cm)**	**Median (cm)**	**Repeat median (cm)**	**Cases *n***	**RR (95% CI)[a]**	**RR (95% CI)[b]**	**RR (95% CI)[c]**	**RR (95% CI)[c]**	**RR (95% CI)[d]**
1	48-76	72	74	576	1	1	1	**Per SD increase**	**Per SD increase**
2	76.1-84	81	83	757	1.23 (1.10-1.37)	1.22 (1.09-1.36)	1.23 (1.10-1.37)	1.23 (1.17-1.29)	1.07 (0.99-1.16)
3	84.1-93	89	90	750	1.33 (1.19-1.48)	1.31 (1.18-1.47)	1.33 (1.19-1.49)		
4	93.1-171	101	100	830	1.52 (1.37-1.69)	1.51 (1.36-1.69)	1.57 (1.41-1.76)		
χ^2								246.3	
Hip quartiles	**Range (cm)**	**Median (cm)**	**Repeat median (cm)**	**Cases *n***	**RR (95% CI)[a]**	**RR (95% CI)[b]**	**RR (95% CI)[c]**	**RR (95% CI)[c]**	**RR (95% CI)[d]**
1	62.0-97.0	94	95	658	1	1	1	**Per SD increase**	**Per SD increase**
2	97.1-102.0	100	101	693	1.31 (1.17-1.45)	1.28 (1.15-1.42)	1.29 (1.16-1.44)	1.21 (1.15-1.27)	1.01 (0.92-1.10)
3	102.1-109	105	106	772	1.35 (1.22-1.50)	1.32 (1.19-1.47)	1.34 (1.21-1.50)		
4	109.1-179	116	116	790	1.50 (1.35-1.66)	1.47 (1.32-1.63)	1.53 (1.38-1.71)		
χ^2								235.8	
WHR quartiles	**Range**	**Median**	**Repeat median**	**Cases *n***	**RR (95% CI)[a]**	**RR (95% CI)[b]**	**RR (95% CI)[c]**	**RR (95% CI)[c]**	**RR (95% CI)[d]**
1	0.503-0.773	0.745	0.771	673	1	1	1	**Per SD increase**	**Per SD increase**
2	0.775-0.820	0.798	0.817	680	0.99 (0.89-1.11)	1.0 (0.90-1.11)	1.00 (0.90-1.11)	1.14 (1.08-1.21)	1.03 (0.97-1.10)
3	0.820-0.870	0.843	0.848	718	1.03 (0.93-1.15)	1.04 (0.94-1.16)	1.05 (0.94-1.17)		
4	0.870-1.351	0.905	0.889	842	1.23 (1.11-1.37)	1.25 (1.13-1.39)	1.27 (1.14-1.41)		
χ^2								202.0	

Table continued

Trunk fat mass quartiles	Range (kg)	Median (kg)	Repeat median (kg)	Cases *n*	RR (95% CI)[a]	RR (95% CI)[b]	RR (95% CI)[c]	RR (95% CI)[c]
1	0.6-10.2	8.4	8.6	545	1	1	1	**Per SD increase**
2	10.3-13.3	11.8	12	731	1.29 (1.15-1.44)	1.27 (1.13-1.42)	1.28 (1.14-1.43)	1.25 (1.19-1.31)
3	13.4-16.9	15	14.8	741	1.33 (1.19-1.48)	1.30 (1.16-1.46)	1.33 (1.18-1.49)	
4	17.0-48.3	19.9	19.3	896	1.66 (1.49-1.84)	1.62 (1.45-1.81)	1.69 (1.51-1.89)	
χ^2								259.4
Trunk fat % quartiles	**Range (%)**	**Median (%)**	**Repeat median (%)**	**Cases *n***	**RR (95% CI)[a]**	**RR (95% CI)[b]**	**RR (95% CI)[c]**	**RR (95% CI)[c]**
1	3.0-30.0	26.2	27.2	540	1	1	1	**Per SD increase**
2	30.1-35.2	32.8	33.5	739	1.33 (1.19-1.48)	1.31 (1.17-1.47)	1.32 (1.18-1.47)	1.25 (1.18-1.31)
3	35.3-40.0	37.5	37.6	753	1.36 (1.22-1.52)	1.34 (1.20-1.50)	1.36 (1.21-1.52)	
4	40.1-75.6	43.3	42.5	881	1.61 (1.45-1.80)	1.59 (1.42-1.77)	1.63 (1.46-1.83)	
χ^2								250.9

[a] stratified by age at recruitment, region of recruitment, and socioeconomic status (Townsend deprivation index)

[b] adjusted for family history of breast cancer, age at menarche, age at first birth, parity, age at menopause, previous HRT use, smoking, alcohol intake frequency, physical activity, height, and ethnicity

[c] further adjusted for body size at age 10

[d] further adjusted for body fat mass

Table 8.5: Linear trend for adiposity and breast cancer risk using repeat assessment of baseline measures by impedance compared to DXA measures

	Using baseline measures	Using DXA measures
Body fat mass, HR (95% CI) per 10 kg increase	1.25 (1.19-1.31)	1.27 (1.20-1.33)
Body fat percentage, HR (95% CI) per 10% increase	1.34 (1.25-1.44)	1.36 (1.26-1.46)
Trunk fat mass, HR (95% CI) per 10 kg increase	1.52 (1.39-1.67)	1.52 (1.39-1.67)
Trunk fat percentage, HR (95% CI) per 10% increase	1.31 (1.23-1.39)	1.30 (1.22-1.38)

All analyses stratified by age at recruitment, region of recruitment, and Townsend deprivation index
All analyses adjusted for adjusted for family history of breast cancer, age at menarche, age at first birth, parity, age at menopause, previous HRT use, smoking, alcohol intake frequency, physical activity, height, ethnicity, and body size at age 10

Table 8.6: Association of body size/composition indices with invasive breast cancer risk in women who have been postmenopausal for less than 12 years versus 12 or more years

	<12 years since menopause			≥12 years since menopause		
	Quartile	**Cases, *n***	**RR (95% CI)**	**Quartile**	**Cases, *n***	**RR (95% CI)**
Weight	1	288	1	1	283	1
	2	324	1.18 (0.99-1.40)	2	331	1.31 (1.10-1.55)
	3	284	1.05 (0.88-1.26)	3	387	1.56 (1.32-1.85)
	4	339	1.35 (1.14-1.61)	4	386	1.80 (1.52-2.13)
			p for interaction = 0.0049			
Body mass index	1	332	1	1	273	1
	2	317	1.10 (0.93-1.30)	2	357	1.40 (1.18-1.66)
	3	298	1.18 (1.00-1.41)	3	373	1.49 (1.26-1.77)
	4	298	1.28 (1.08-1.53)	4	383	1.83 (1.54-2.17)
			p for interaction = 0.0702			
Body fat mass	1	292	1	1	255	1
	2	326	1.25 (1.06-1.49)	2	336	1.39 (1.16-1.66)
	3	266	1.08 (0.90-1.29)	3	388	1.66 (1.40-1.97)
	4	337	1.47 (1.24-1.75)	4	383	1.86 (1.56-2.21)
			p for interaction = 0.0009			
Body fat percentage	1	309	1	1	255	1
	2	315	1.19 (1.01-1.41)	2	336	1.36 (1.14-1.62)
	3	290	1.18 (0.99-1.40)	3	373	1.50 (1.26-1.79)
	4	307	1.36 (1.15-1.62)	4	398	1.79 (1.51-2.13)
			p for interaction = 0.1106			
Waist	1	322	1	1	251	1
	2	340	1.16 (0.99-1.37)	2	373	1.37 (1.15-1.63)
	3	278	1.13 (0.95-1.35)	3	390	1.64 (1.38-1.94)
	4	296	1.32 (1.11-1.57)	4	374	1.83 (1.54-2.18)
			p for interaction = 0.0108			
Hip	1	336	1	1	312	1
	2	299	1.17 (0.99-1.38)	2	342	1.49 (1.26-1.76)
	3	305	1.13 (0.96-1.34)	3	385	1.62 (1.37-1.90)
	4	296	1.29 (1.09-1.53)	4	349	1.73 (1.46-2.04)
			p for interaction = 0.0167			
Waist-to-hip ratio	1	346	1	1	300	1
	2	314	1.02 (0.86-1.20)	2	325	1.04 (0.88-1.23)
	3	277	0.97 (0.81-1.14)	3	350	1.12 (0.95-1.33)
	4	299	1.17 (0.99-1.38)	4	413	1.36 (1.16-1.60)
			p for interaction = 0.4025			
Trunk fat mass	1	295	1	1	254	1
	2	319	1.20 (1.01-1.43)	2	348	1.37 (1.15-1.64)
	3	269	1.08 (0.90-1.29)	3	379	1.61 (1.35-1.91)
	4	336	1.43 (1.21-1.70)	4	381	1.87 (1.57-2.23)
			p for interaction = 0.0094			
Trunk fat percentage	1	294	1	1	255	1
	2	320	1.25 (1.05-1.48)	2	346	1.37 (1.14-1.63)
	3	293	1.22 (1.02-1.45)	3	361	1.51 (1.27-1.80)
	4	313	1.42 (1.19-1.69)	4	400	1.82 (1.53-2.16)
			p for interaction = 0.1580			

All analyses adjusted for family history of breast cancer, age at menarche, age at first birth, parity, age at menopause, previous HRT use, smoking, alcohol intake frequency, physical activity, height, ethnicity and body size at age 10

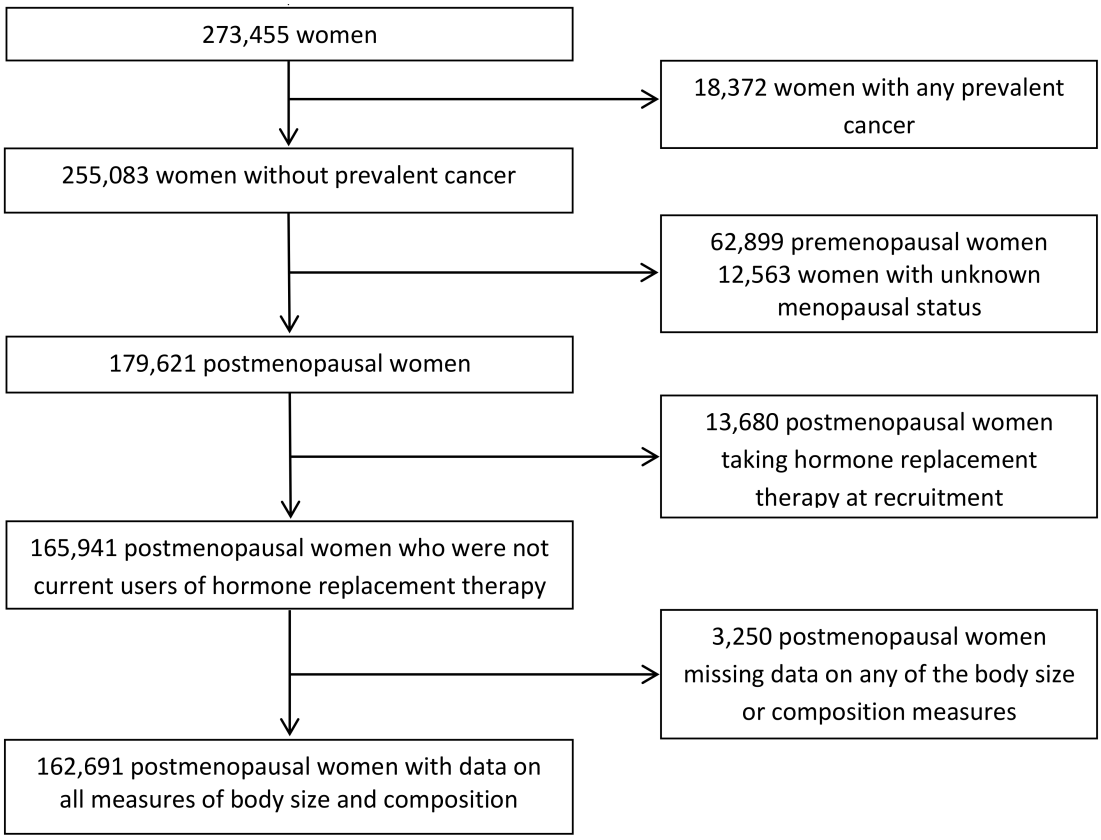
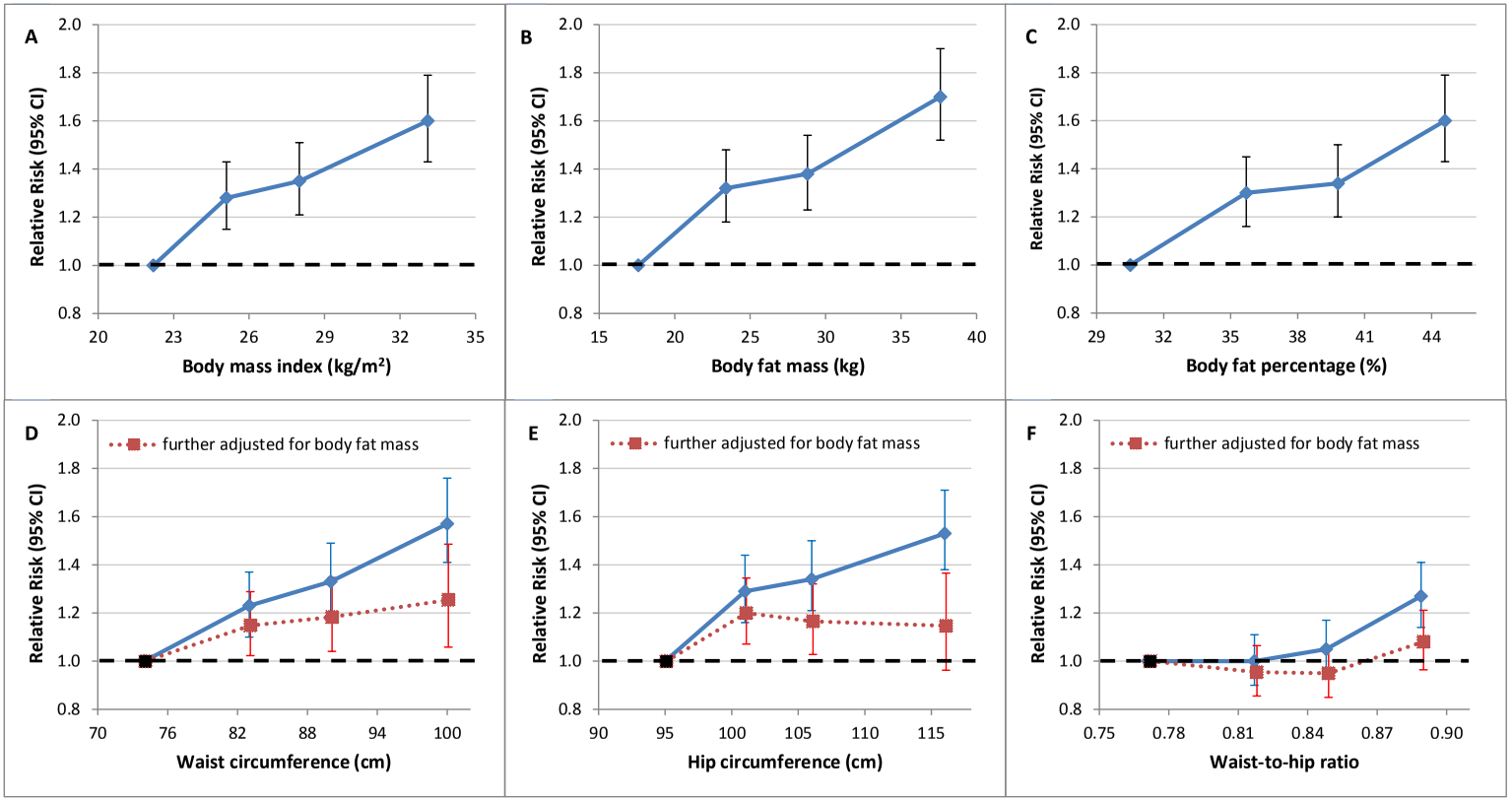


Figure 8.1: Participant inclusion/exclusion flow diagram for analyses of adiposity and breast cancer risk in postmenopausal women



For panels D, E, and F, the blue diamonds represent multivariable adjusted estimates and the red squares represent estimates further adjusted for body fat mass.
Body size/composition measures were grouped into quartiles.
The figure shows point estimates and 95% confidence intervals

Figure 8.2: Association between various measures of adiposity and postmenopausal breast cancer risk: (A) Body mass index, (B) Body fat mass, (C) Body fat percentage, (D) Waist circumference, (E) Hip circumference, (F) Waist-to-hip ratio

Chapter 9

PHYSICAL ACTIVITY AND BREAST CANCER RISK

9.1 Introduction

Previous prospective studies have assessed the association between self-reported physical activity and breast cancer, with overall findings that support a protective role for physical activity in breast cancer risk[44,223,224]. However, self-reported responses from questionnaires are prone to both inaccurate reporting and reporting bias, leading to random and systematic error, respectively[9]. Further research incorporating objective methods of physical activity assessment, such as accelerometers, is needed to address this measurement error and reduce misclassification of physical activity levels.

Some studies have also reported greater reductions in risk associated with physical activity among postmenopausal women[223] and those with a relatively low body mass index (BMI)[225] but it is unclear as to what extent these observed differences

may be attributed to differential reporting bias in self-reported estimates of physical activity by subgroups of menopausal status and BMI, as opposed to genuine effect modification.

This chapter evaluates the associations between physical activity and breast cancer risk in over 47,000 pre- and 126,000 postmenopausal women, using baseline questionnaire-based physical activity in conjunction with subsequently measured accelerometer values. I also examine whether and to what extent the relationship between physical activity and breast cancer risk may be mediated by adiposity using an objective measure of body fat assessed by bioelectrical impedance, rather than self-reported BMI used by many other cohorts. Impedance measures of body fat showed a marginally stronger association with breast cancer risk compared to BMI, as demonstrated in Chapter 8.

9.2 Methods

9.2.1 Data source

Details are provided in previous chapters.

9.2.2 Study participants

The complete UK Biobank dataset at the time of analysis for the this chapter included 502,620 UK adults (229,165 men and 273,455 women) aged between 40-69 years at recruitment during 2006-2010.

Women were excluded from the analyses if they had a prior cancer diagnosis (except for non-melanoma skin cancer ICD-10 C44) (n=18,372), had missing data

on any of the physical activity variables used in the analyses (n=65,291), had a sum of all walking, moderate and vigorous physical activity greater than 6,720 minutes per week as detailed in the International Physical Activity Questionnaire (IPAQ) guidelines (n=2,021)[135], and had missing data on body fat mass (n=3,163). A total of 184,608 women were included in the final analyses.

Women who had unknown menopausal status were excluded from the subgroup analysis by menopausal status (n=10,448). In the final subgroup analyses by menopausal status, 47,456 premenopausal and 126,704 postmenopausal women were included (Figure 9.1). Women were defined as being premenopausal or postmenopausal at recruitment based on whether they reported their periods had stopped; for these analyses, women with unknown self-reported menopausal status who were under the age of 45 who had not undergone a bilateral oophorectomy were categorised as premenopausal and women age 53 or over and/or had had both ovaries removed were categorised as postmenopausal, based on previously established criteria[84].

9.2.3 Physical activity

Physical activity questions from the baseline questionnaire captured the frequency and duration of three levels of activity (walking, moderate, and vigorous). Further details on self-reported physical activity from the baseline questionnaire are detailed in the methods section of Chapter 6 (section 6.2.3). Questions from the baseline questionnaire are shown in Appendix A.

Between February 2013 and December 2015, 236,519 participants, all of whom had provided a valid email address, were invited to participate in a seven day

accelerometer study. Accelerometer data was available on a total of 40,984 women in the final analyses. Details on the accelerometer measures can be found in Chapter 6 (section 6.2.4). Further information on data collection, processing, and analysis can be found elsewhere[111].

9.2.4 Anthropometry and body composition

UK Biobank derived BMI by dividing weight (kg) by the square of standing height (m^2). At the baseline interview, the Tanita BC-418MA body composition analyser was used to measure body fat mass using bio-impedance.

9.2.5 Ascertainment of cancer cases

For details on ascertainment of cancer cases, as well as a discussion on limitations of the flagging of the cancer registry and other data quality considerations, please refer to Chapter 8, section 2.4.

9.2.6 Statistical analyses

Baseline characteristics of participants were summarised by physical activity separately for pre- and postmenopausal women. Women were followed from date of baseline assessment centre visit until the earliest of: date of breast cancer registration (ICD-10 C50), date of death, date of loss to follow-up or end of follow-up for cancer incidence (30 November 2014). Women diagnosed with any cancer other than breast cancer (with the exception of non-melanoma skin cancer) during follow-up were censored at date of diagnosis.

Multivariable-adjusted Cox regression with attained age as the underlying time

variable was used to estimate hazard ratios (referred to as relative risks [RRs]) and 95% confidence intervals (CIs) for the associations between physical activity and breast cancer risk. Relative risks are presented with group-specific confidence intervals (95% g-s CI) for the log risk in each group, which allows comparisons to be made between any two categories, even if neither is the reference group[226]. Conventional 95% CIs are reported in the text.

For the risk analyses, in which separate models were used for premenopausal and postmenopausal women, physical activity was categorised into fourths based on separate distributions in pre- and postmenopausal women. Assessment of interaction terms between each exposure of interest and the underlying time variable did not suggest significant deviation from proportional hazards. In order to use values that are representative of objective measures of physical activity, tests for linear trend were performed with categories coded in an ordinal fashion using the accelerometer-based physical activity median within each baseline physical activity category as a surrogate for questionnaire-based activity[205]. Participants with poor wear time (n=2,815) or poor data calibration (n=4) due to insufficient data were excluded when recoding categories according to the accelerometer medians. Linear trends are reported per 5 milli-gravity of accelerometer-based physical activity. However, it is important to note that these two methods of measuring physical activity are not directly comparable given that they have unique strengths and limitations. For example, although the wrist accelerometers used in the present study are able to capture physical activity objectively 24 hours a day, they are unable to reliably capture certain activities such as cycling[158].

Covariates were selected a priori based on known risk factors for breast cancer. All analyses were stratified by 5-year age at recruitment categories, region

of recruitment, and socioeconomic status (based on fifths of Townsend deprivation index[119]), and adjusted for family history of breast cancer (no, yes), height ($<$160, 160-165.9, $\geq$166 cm), number of births (nulliparous, 1-2, $\geq$3 births), age at first birth ($<$25, 25-29, $\geq$30 years), age at menarche ($<$12, 12-13, $\geq$14 years), oral contraceptive use (never, previous, current), ever had breast cancer screening (no, yes), smoking status (never, previous, current), and alcohol intake frequency (less than 3 times a month, 1-4 times a week, daily or almost daily). Analyses for postmenopausal women were additionally adjusted for hormone replacement therapy (HRT) use (never, previous, current) and age at menopause ($<$45, 45-49, 50-54, $\geq$55 years). Analyses were further adjusted for body fat mass in multivariable adjusted models that included height. In models which included all women, the association of body fat mass with risk was allowed to vary by menopausal status by including a variable generated by the cross-classification of body fat mass and menopausal status.

Women with missing values for any of the adjustment variables were assigned to a separate "unknown" category for the respective variable. Information was either missing or reported as unknown for $<$3% of covariates, with the exception of age at menopause (19.3%) among postmenopausal women, which may be difficult to accurately report since menopausal status can potentially be masked by hysterectomy or HRT use before menstrual periods stop naturally[84]. A sensitivity analysis restricted to participants with known values for all adjustment variables was conducted to assess the impact of missing values. I also conducted sensitivity analyses to assess the impact of excluding the first two years of follow-up, excluding those who reported long-term illness, disability, or infirmity, excluding those who reported poor health, and excluding follow-up for women once they reached

age 50 in the premenopausal analysis.

A chi-squared test for heterogeneity was used to assess whether menopausal status was an effect modifier in the associations between accelerometer-based physical activity and breast cancer risk. Before examining whether there was heterogeneity by BMI, I assessed whether reporting of amount of physical activity varied by BMI by comparing baseline questionnaire-based to accelerometer-based physical activity across BMI groups. It should be noted that this analysis cannot capture random and systematic measurement error in isolation. Instead, the results are likely reflective of a combination of different types of measurement error and biases. Questionnaire-based activity was first categorised into eighths based on the overall study population. The accelerometer-based median of physical activity was then plotted against the median questionnaire-based physical activity value within each eighth, separately in women with BMI $<$25 kg/m^2, 25-29.9 kg/m^2, and $\geq$30 kg/m^2.

Chi-squared tests for heterogeneity were also used to determine whether the association between accelerometer-based physical activity and breast cancer risk was modified by BMI group ($<$25 kg/m^2, $\geq$25 kg/m^2), number of births (0-1, $\geq$2 births), HRT use (never, ever), and alcohol intake (less or more than one drink weekly). All analyses were conducted using STATA version 15.0 (Stata Corp LP, College Station, TX).

9.3 Results

A total of 3,189 cases were diagnosed among 184,608 women during a mean follow-up of 5.7 (SD 1.1) years. Accelerometer data were available on a subset of 40,984

women in these analyses. For the subgroup analyses by menopausal status, 717 cases were diagnosed among 47,456 premenopausal women and 2,315 invasive breast cancer cases were diagnosed among 126,704 postmenopausal women. I examined physical activity in relation to potential confounders for breast cancer (Table 9.1). Women who did the least physical activity had a higher BMI, greater body fat mass and were slightly more likely to be smokers. A slightly greater proportion of women in the top fourth of physical activity consumed alcohol on a weekly or more frequent basis. Compared to premenopausal women, postmenopausal women reported more physical activity but were objectively less physically active, as measured by accelerometer.

Self-reported physical activity was associated with breast cancer risk reduction in both premenopausal and postmenopausal women (Table 9.2). Premenopausal women in the top fourth of physical activity, who reported 53.6 or more MET-hours/week of physical activity, had a 19% decreased risk of breast cancer (RR 0.81; 95% CI 0.65-1.00) compared to those in the bottom fourth. The protective association was strengthened after further adjusting for body fat mass (RR 0.79; 95% CI 0.63-0.98). Among postmenopausal women, the top fourth of physical activity (59.9 or more MET-hours/week), was associated with a 17% decreased risk of breast cancer (RR 0.83; 95% CI 0.74-0.93) compared to the bottom fourth. Further adjustment for body fat mass attenuated this protective association (RR 0.88; 95% CI 0.78-0.99).

The risk associated with accelerometer-based physical activity, per 5 milligravity, was similar in premenopausal women (RR 0.81; 95% CI 0.67-0.97) and postmenopausal women (RR 0.78; 95% CI 0.67-0.90). After adjustment for body fat mass, accelerometer-based physical activity was associated with a 21% de-

creased risk of breast cancer in premenopausal women (RR 0.79; 95% CI 0.65-0.95) and a 16% decreased risk of breast cancer in postmenopausal women (RR 0.84; 95% CI 0.72-0.97). There was no heterogeneity in accelerometer-based trends by menopausal status (p=0.62) (Table 9.2). Figure 9.2 illustrates the association between physical activity and breast cancer risk for all women combined. The results of sensitivity analyses did not differ materially from the main findings. Table 9.4 shows the results of sensitivity analyses for the association between self-reported physical activity and breast cancer risk in all women, excluding the first two years of follow-up.

In order to assess potential differences in measurement error by BMI, accelerometer medians within each eighth of questionnaire-based physical activity were compared across BMI groups. For each level of questionnaire-based physical activity, differences in objective accelerometer-based physical activity medians were seen across the three BMI groups (Figure 9.3). The range in the median values of accelerometer-based physical activity for women in the bottom compared to the top level of questionnaire-based physical activity (in milli-gravity) was 26.9-32.0 in women with BMI$<$25 kg/m^2, 25.1-29.0 in women with BMI 25-29.9 kg/m^2, and 22.9-25.7 in women with BMI of 30 kg/m^2 or more.

The inverse association between accelerometer-based physical activity and breast cancer risk in both women with BMI$<$25 kg/m^2 and BMI$\geq$25 kg/m^2 was of similar magnitude per 5 milli-gravity (RR 0.80; 95% CI 0.67-0.95; p for trend=0.011 and RR 0.82; 95% CI 0.68-0.99; p for trend=0.036, respectively; p for heterogeneity=0.82) (Table 9.3); nor was there evidence of heterogeneity in accelerometer-based physical activity by HRT use in postmenopausal women (p=0.12), number of births (p=0.85), or alcohol intake frequency (p=0.70) (Figure 9.4).

9.4 Discussion

In this large prospective study, physical activity was associated with an approximately 19% reduced risk of breast cancer in premenopausal women and 22% reduced risk in postmenopausal women per 5 milli-gravity of accelerometer-based physical activity, before adjustment for body fat mass. Adjusting for body fat mass slightly strengthened the inverse association in premenopausal women but attenuated the association in postmenopausal women, and the adiposity-adjusted trends in risk with increasing accelerometer-based physical activity did not vary significantly by menopausal status. These findings suggest a protective effect of physical activity on breast cancer risk beyond the role of adiposity.

In postmenopausal women, the finding that adjusting for body fat mass attenuated the protective association between physical activity and breast cancer is consistent with results from a previous postmenopausal cohort[227], although other cohort studies in postmenopausal women found minimal change after adjusting for self-reported BMI[228,229]. In the current analyses, adjusting for direct measures of body fat mass rather than self-reported BMI may explain the greater magnitude of attenuation compared to those reported in prior studies. The fact that the protective association of physical activity with postmenopausal breast cancer risk was somewhat attenuated after adjusting for body fat mass supports the notion that estrogen-dependent mechanisms are involved. Since high BMI is associated with increased breast cancer risk in postmenopausal women[230], the association of physical activity with breast cancer risk observed before adjustment for body fat mass is likely to be partly mediated by adiposity. Higher levels of body fat have been hypothesized to increase postmenopausal breast cancer risk

through estrogen-stimulated carcinogenesis[11,201] and lower concentrations of sex hormone-binding globulin, resulting in increased estrogen bioavailability[22].

In contrast to the association in postmenopausal women, high BMI is generally inversely associated with breast cancer risk in premenopausal women[230] and, as expected, adjusting for adiposity strengthened the protective association between physical activity and breast cancer risk. Therefore the protective association between physical activity and breast cancer risk in premenopausal women is unlikely to be due to residual confounding by BMI. The mechanisms behind the association between physical activity and breast cancer risk reduction observed after adjustment for adiposity remain unclear, but may be due to a combination of effects involving pathways such as improved insulin sensitivity, reduced chronic inflammation, and enhanced immune function[10,47,48].

A recent meta-analysis of 18 cohort and 11 case-control studies, all based on self-reported physical activity, found physical activity was associated with the greatest magnitude reduction in breast cancer risk among women with lower BMI and there was no significant association between physical activity and breast cancer risk among women who were obese[225]. One factor that might partly explain the lack of physical activity associated risk reduction is over-reporting of physical activity among individuals who are obese[231]. In these data, women with BMI$<$25 kg/m^2 who were in the least active category of questionnaire-based activity did more accelerometer-based physical activity, on average, than women with BMI$\geq$30 kg/m^2 whose questionnaire-based physical activity categorised them into the most active group, suggesting substantial over-reporting of physical activity in women with higher BMI. After examination of accelerometer-based measures in relation to breast cancer risk, which should eliminate the physical activity reporting bias

that may vary by BMI, there was no heterogeneity by BMI group and the tests for linear trend showed similar magnitudes of association. These results are consistent with those from a recent study that pooled data from 10 prospective cohorts, all based on self-reported estimates of physical activity from questionnaire data as the exposure, with incident breast cancer as an outcome[224].

Given that questionnaire-based physical activity alone is poorly reported, especially in certain subgroups[232,233], a notable strength of the present study is the availability of objectively-measured accelerometer-based physical activity. While only questionnaire-based physical activity was available on the whole cohort at baseline, using accelerometer-based physical activity data as a surrogate mitigates the measurement error associated with questionnaire-based physical activity. Since self-reported measures of adiposity such as weight tend to be underestimated among overweight and obese women[202,234], another strength of the current study is the availability of objectively-measured body fat. Other strengths include the prospective design, large sample size, virtually complete follow-up[127], and the availability of data on a wide range of potential confounders. In the present study, I chose to take a traditional approach, comparable to those utilised by recent studies, to account for the role of adiposity as a mediator in the relationship between physical activity and breast cancer risk. Although I recognize that there are more sophisticated, emerging methods for mediation analysis that are being adopted by researchers with extensive statistical backgrounds, the implementation of these methods may be complex and must satisfy strict assumptions in order to produce results that are interpretable and valid[235]. For the majority of women during the follow-up period, updated data was not available on menopausal status, which may bias the subgroup analyses by menopausal status. Another limitation of the

study is the lack of information on tumor characteristics such as stage, grade, and hormone receptor status, although these data will be available in UK Biobank in the future.

9.5 Conclusion

In conclusion, the results presented in this chapter demonstrate a protective association between physical activity and breast cancer risk in both pre- and postmenopausal women. After adjusting for body fat mass, a similar magnitude of inverse association between physical activity and breast cancer risk remained in both groups and may be mediated by non-adiposity-related factors. Future research should aim to further clarify the mechanisms underlying this reduction in risk associated with physical activity.

Table 9.1: Characteristics of the study population at recruitment, according to self-reported physical activity and menopausal status

	Premenopausal		Postmenopausal	
Fourths of physical activity	**Bottom fourth (least active)**	**Top fourth (most active)**	**Bottom fourth (least active)**	**Top fourth (most active)**
Age at recruitment (years), mean (SD)	45.7 (3.6)	45.6 (3.5)	59.3 (5.5)	60.4 (5.5)
Lowest fifth of socioeconomic status, %	21.0	22.7	18.3	17.9
Family history of breast cancer, %	5.7	5.7	6.6	6.6
Age at menarche (years), mean (SD)	13.0 (1.6)	13.1 (1.6)	12.9 (1.6)	13.0 (1.7)
Age at first birth (years), mean (SD)	28.2 (5.6)	27.4 (5.4)	25.7 (5.0)	25.2 (4.8)
Parity, mean (SD)	1.5 (1.2)	1.6 (1.2)	1.8 (1.2)	1.9 (1.2)
Ever oral contraceptive use, %	89.3	89.3	80.8	77.8
Age at menopause (years), mean (SD)			49.7 (5.0)	49.7 (5.2)
Current hormone replacement therapy use, %			8.3	7.8
Current smokers, %	11.0	10.8	9.3	7.8
One or more drinks weekly, %	63.1	64.8	61.0	62.7
Height (cm), mean (SD)	164.1 (6.3)	164.2 (6.2)	162.1 (6.3)	161.9 (6.1)
Body mass index (kg/m^2), mean (SD)	27.5 (5.9)	25.6 (4.7)	28.2 (5.7)	26.3 (4.5)
Body fat mass (kg), mean (SD)	28.0 (11.6)	23.6 (9.5)	29.5 (10.7)	25.3 (8.8)
Plumper than average at age 10, %	20.0	18.0	18.9	15.8
Total physical activity (MET-hrs/week), median (IQR)	6.9 (3.5-10.0)	85.7 (65.9-125.7)	7.6 (4.0-10.8)	97.3 (74.7-139.5)
Accelerometer physical activity (milli-gravity), mean (SD)*	28.1 (7.3)	34.2 (9.5)	25.5 (6.7)	29.8 (8.3)

Abbreviations: interquartile range (IQR), metabolic equivalent hours per week (MET-hrs/week), standard deviation (SD)

*data was collected after recruitment between February 2013 and December 2015

Table 9.2: Association between self-reported and accelerometer-based physical activity and risk of invasive breast cancer by menopausal status

Self-reported physical activity quartiles (MET-hrs/week)	Median self-reported physical activity (MET-hrs/week)	Repeat median self-reported physical activity (MET-hrs/week)	Median accelerometer physical activity (milli-gravity/week)	Cases (n)	Relative Risk (95% g-s CI)[a]	Relative Risk (95% g-s CI)[b]
Premenopausal women						
0-13.3	7.0	15.0	27.3	198	1.00 (0.87-1.15)	1.00 (0.87-1.15)
13.4-28.2	20.2	22.8	29.2	198	0.99 (0.86-1.14)	0.98 (0.85-1.12)
28.3-53.5	38.4	33.9	30.8	168	0.85 (0.73-0.99)	0.83 (0.71-0.96)
≥53.6	86.0	61.2	32.8	153	0.81 (0.69-0.95)	0.79 (0.67-0.92)
Linear trend per 5 milli-gravity (corrected with accelerometer values)					0.81 (0.67-0.97)	0.79 (0.65-0.95)
P for trend					0.023	0.011
Postmenopausal women						
0-13.9	7.7	13.3	25.0	621	1.00 (0.92-1.08)	1.00 (0.92-1.08)
13.96-30.1	21.4	24.0	26.7	612	0.97 (0.90-1.06)	1.00 (0.93-1.08)
30.2-59.8	41.9	39.5	27.6	563	0.88 (0.81-0.96)	0.92 (0.84-1.00)
≥59.9	97.3	64.1	28.9	519	0.83 (0.76-0.91)	0.88 (0.80-0.96)
Linear trend per 5 milli-gravity (corrected with accelerometer values)					0.78 (0.67-0.90)	0.84 (0.72-0.97)
P for trend					0.001	0.017
P for heterogeneity by menopausal status (corrected with accelerometer values)= 0.62						

[a] Stratified by age at recruitment, region of recruitment, and socioeconomic status. Adjusted for family history of breast cancer, height, parity, age at menarche, oral contraceptive use, age at first birth, smoking status, ever had breast cancer screening, alcohol intake frequency, age at menopause (only in postmenopausal women), and HRT use (only in postmenopausal women)

[b] further adjusted for quartiles of body fat mass within each menopausal status group

Table 9.3: Association between self-reported and accelerometer-based physical activity and risk of invasive breast cancer by body mass index

Self-reported physical activity quartiles (MET-hrs/week)	Median self-reported physical activity (MET-hrs/week)	Repeat median self-reported physical activity (MET-hrs/week)	Median accelerometer physical activity (milli-gravity/week)	Cases (n)	Relative Risk (95% g-s CI)[a]	Relative Risk (95% g-s CI)[b]
Women with body mass index <25						
<13.6	7.8	14.6	27.3	288	1.00 (0.89-1.12)	1.00 (0.8-1.12)
13.6-29.4	21.0	24.6	29.1	296	0.82 (0.73-0.92)	0.83 (0.74-0.93)
29.5-58.2	41.0	38.1	30.3	310	0.79 (0.71-0.88)	0.80 (0.72-0.90)
≥58.3	94.5	67.6	31.9	312	0.79 (0.71-0.88)	0.81 (0.72-0.91)
Linear trend per 5 milli-gravity (corrected with accelerometer values)					0.78 (0.65-0.93)	0.80 (0.67-0.95)
P for trend					0.005	0.011
Women with body mass index ≥25						
<13.6	6.9	12.6	24.6	571	1.00 (0.92-1.09)	1.00 (0.92-1.09)
13.6-29.4	20.8	21.9	26.1	549	1.07 (0.98-1.16)	1.08 (0.99-1.17)
29.5-58.2	40.6	36.9	27.0	466	0.96 (0.88-1.05)	0.97 (0.89-1.07)
≥58.3	94.2	57.5	27.9	397	0.85 (0.77-0.94)	0.87 (0.79-0.96)
Linear trend per 5 milli-gravity (corrected with accelerometer values)					0.80 (0.67-0.96)	0.82 (0.68-0.99)
P for trend					0.018	0.036
P for heterogeneity by body mass index (corrected with accelerometer values)= 0.82						

[a] Stratified by age at recruitment, region of recruitment, and socioeconomic status. Adjusted for family history of breast cancer, height, parity, age at menarche, oral contraceptive use, age at first birth, smoking status, ever had breast cancer screening, alcohol intake frequency, age at menopause (only in postmenopausal women), and HRT use (only in postmenopausal women)

[b] further adjusted for quartiles of body fat mass within each body mass index group

Table 9.4: Association between self-reported physical activity and risk of invasive breast cancer, excluding the first two years of follow-up

Physical activity quartiles	Cases (n)	Relative Risk (95% g-s CI)[a]
Quartile 1	558	1.00 (0.92-1.09)
Quartile 2	578	1.04 (0.96-1.13)
Quartile 3	538	0.98 (0.90-1.07)
Quartile 4	463	0.87 (0.80-0.96)

[a] Stratified by age at recruitment, region of recruitment, and socioeconomic status. Adjusted for family history of breast cancer, height, parity, age at menarche, oral contraceptive use, age at first birth, smoking status, ever had breast cancer screening, alcohol intake frequency, age at menopause, HRT use, and quartiles of body fat mass within each menopausal status group

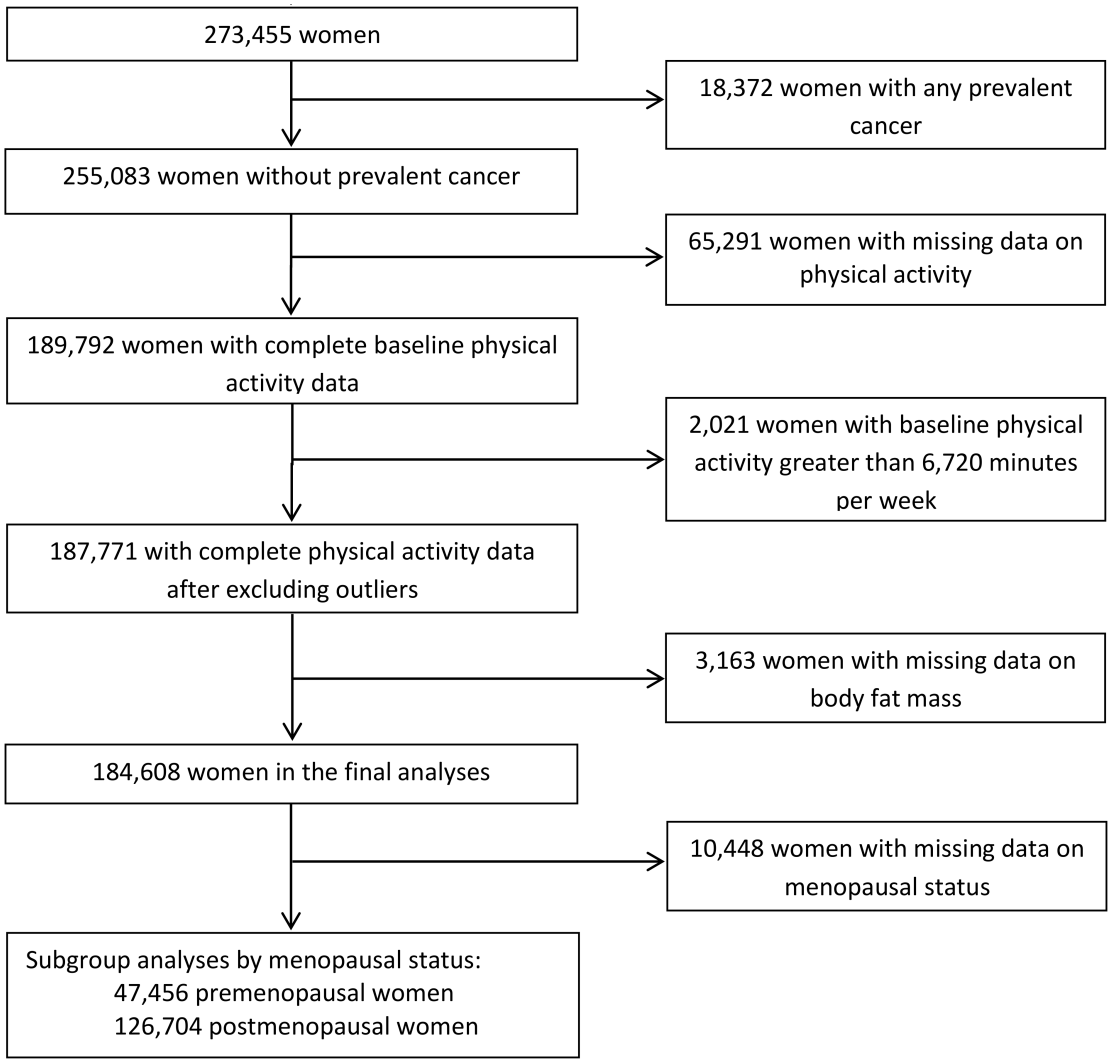
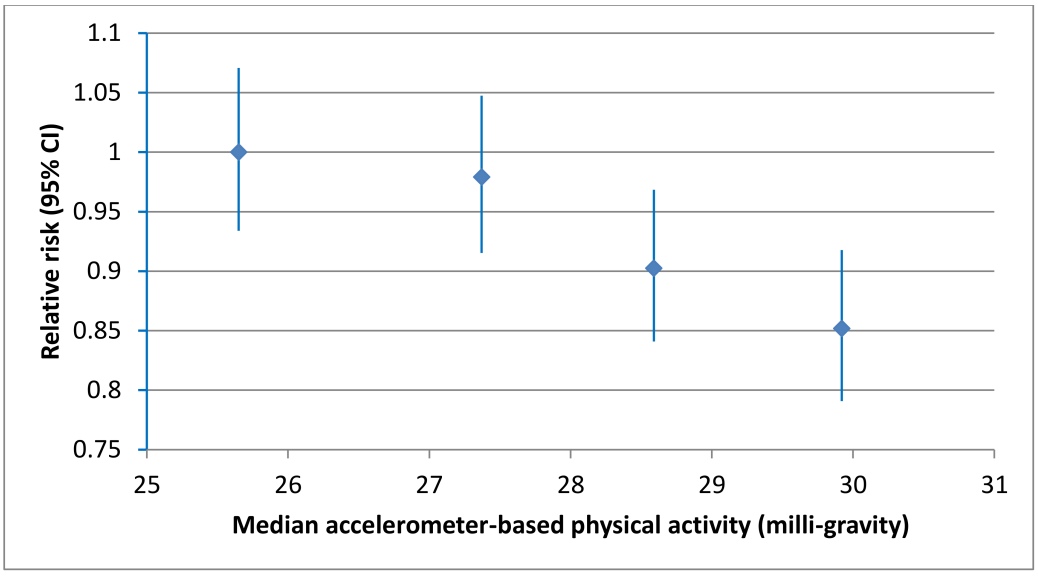
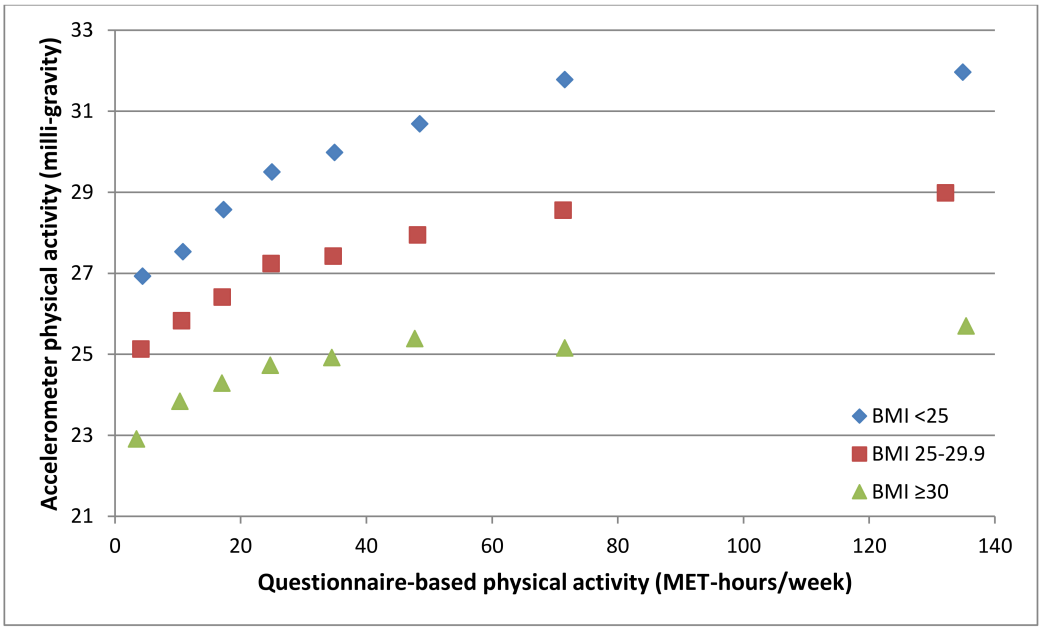


Figure 9.1: Participant inclusion/exclusion flow diagram for analyses of physical activity in relation to breast cancer risk



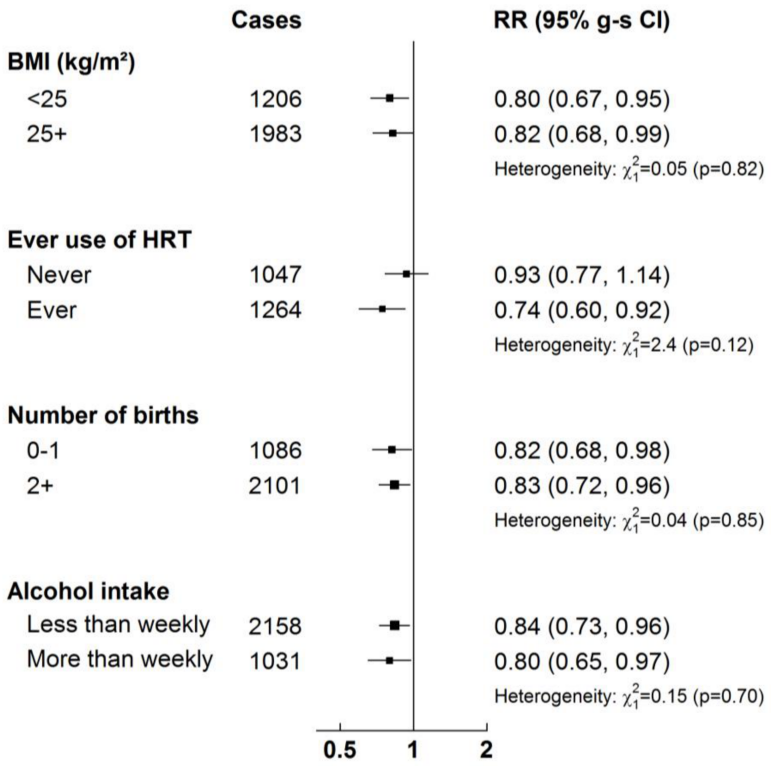
Relative risk for the association between questionnaire-based physical activity and risk of invasive breast cancer are plotted against the median accelerometer-based physical activity value within each fourth of questionnaire-based physical activity. The figure shows point estimates and 95% group-specific confidence intervals.
Abbreviation: CI, confidence interval

Figure 9.2: Association between accelerometer-based physical activity and risk of invasive breast cancer



The blue diamonds represent the accelerometer medians for women with BMI<25kg/m^2.
The red squares represent the accelerometer medians for women with BMI 25-29.9 kg/m^2.
The green triangles represent the accelerometer medians for women with BMI of 30 kg/m^2 or more.
Abbreviations: BMI (body mass index), MET (metabolic equivalent)

Figure 9.3: Accelerometer-based physical activity median within each category of questionnaire-based physical activity, according to BMI (in units of kg/m^2) group



Analyses stratified by ever use of hormone replacement therapy were restricted to postmenopausal women.
The figure shows the linear trend per 5 milli-gravity of accelerometer-based physical activity for each subgroup and 95% group-specific confidence intervals.
Abbreviations: body mass index (BMI), hormone replacement therapy (HRT), relative risk (RR)

Figure 9.4: Relative risk of accelerometer-based physical activity and breast cancer risk by BMI, ever use of HRT, number of births and alcohol intake

Chapter 10

DISCUSSION

10.1 Aims of thesis

My main aims were to investigate risk factors for breast cancer in UK Biobank, with a focus on selected modifiable lifestyle factors, and to better characterise the associations between the risk factors themselves. Specifically, the modifiable risk factors identified for further investigation were body size, body composition, and physical activity.

I investigated the cross-sectional associations of self-reported, questionnaire-based physical activity with various measures of body size and composition and then compared these with corresponding associations using objectively-measured physical activity captured by accelerometer. I also explored whether genetic susceptibility to adiposity, as represented by a weighted genetic risk score, is modified by physical activity. Next, I explored the prospective associations between the various indices of body size and composition and breast cancer risk in postmenopausal women, with the primary objective of assessing whether direct measures of fat mass

are better indicators of risk than indirect measures of adiposity, such as BMI. Finally, I investigated the prospective associations of self-reported physical activity, as well as accelerometer-based indices of physical activity, in relation to breast cancer risk.

10.2 Findings in context

The findings presented in this thesis provide novel insights into the relationships between adiposity, physical activity, and breast cancer risk in premenopausal and postmenopausal women.

10.2.1 Risk factors for breast cancer in UK Biobank

In Chapter 4, I investigated the association of breast cancer risk with various anthropometric, lifestyle, reproductive, and other factors previously identified in the literature to be associated with breast cancer risk but which have not been reported on in the UK Biobank prospective cohort.

Height was associated with increased risk, which is in line with prior studies[18]. For comparative body size at age 10, being plumper than average was protective while increased risk was seen in those who were comparatively thinner, which is consistent with the literature[123,217].

BMI was associated with increased risk in postmenopausal women but the results were not significant in premenopausal women. The evidence for an association between adiposity and premenopausal breast cancer is less consistent than for postmenopausal breast cancer. Due to the small numbers of premenopausal breast cancer cases in the current study, the analyses presented in this thesis are likely

to have been underpowered to detect an inverse association between adiposity and premenopausal breast cancer risk as reported by some previous pooled analyses and a large cohort of premenopausal women[26,38,39,236]. The Premenopausal Breast Cancer Collaboration, a large international collaboration aimed at elucidating risk factors for breast cancer that may differ in premenopausal or perimenopausal compared to postmenopausal women, will be able to better characterise the associations between adiposity, physical activity, and premenopausal breast cancer risk overall and by various factors such as tumour subtype and reproductive factors[27].

The variation in the literature concerning whether adiposity is associated with reduced breast cancer risk in premenopausal women may partly be due to differences in the association by tumour subtype, and/or ethnicity. Among prospective cohort studies, higher BMI has been associated with reduced risk of ER^+PR^+ premenopausal breast cancer[31–33]. In contrast, no association between adiposity and hormone receptor negative premenopausal breast cancer was found in the European Prospective Investigation into Cancer and Nutrition[31] and greater waist circumference was associated with increased risk of ER^- PR^- and triple negative premenopausal breast cancer in the Nurses' Health Study II[33]. The association between adiposity and premenopausal breast cancer has also been observed to vary by ethnicity. In a meta-analysis that performed stratified analyses by ethnicity, the inverse association between adiposity and premenopausal breast cancer risk remained significant among African and Caucasian women, but there was a significant positive association observed for Asian women[237]. In a recent prospective study of African American women, central adiposity was associated with increased risk of early-onset (defined as diagnosis before the age of 45 years) ER^- breast cancer[238].

For lifestyle factors, greater physical activity was associated with lower risk and greater alcohol intake was associated with increased risk, consistent with prior studies[44,58]. A borderline significant increased risk was seen in current smokers, although there may be stronger associations within certain subgroups as shown by a recent, large pooled analysis[65].

Reproductive factors are among the most established risk factors for breast cancer[2]. Older age at menarche and a combination of younger age at first birth and higher parity were protective. Current use of oral contraceptives in premenopausal women, older age at menopause, and current use of hormone replacement therapy in postmenopausal women were associated with increased risk. Having a family history of breast cancer was associated with increased risk.

Overall, the patterns of association for established breast cancer risk factors in UK Biobank are consistent with the literature. The results reported in this chapter reflect positively on the high quality of the data collected in this large, prospective cohort and provide a solid basis for the more in-depth investigation of adiposity and physical activity in relation to breast cancer risk in later chapters.

10.2.2 Self-reported physical activity in relation to body size and composition

The cross-sectional associations of self-reported physical activity in relation to various indices of body size and composition were examined in Chapter 5. Specifically, BMI, body fat percentage, body fat mass, trunk fat percentage, trunk fat mass, waist circumference, hip circumference, and waist-to-hip ratio were all similarly lower in women with higher physical activity. These results are in line with

the findings reported from a large prospective cohort study in China with bioimpedance measures[239]. All indices of total and central adiposity were highly correlated, with the exception of waist-to-hip ratio. This suggests that, at the population level, BMI is a fairly good indicator of body size and composition.

Within each category of self-reported total physical activity, a larger proportion of vigorous physical activity was associated with lower adiposity. The results for vigorous physical activity suggest that exercise of higher intensity may be associated with lower adiposity beyond the role of frequency and duration of physical activity, which has been suggested by some prior studies[143]. However, given that the self-reported nature of the physical activity data used in this analysis is a major limitation, it is likely that this result is at least partly due to measurement error. The interpretation of analyses using self-reported versus more objective measures of physical activity was explored further in the next chapter. However, further research would be needed to establish reliable intensity thresholds for accelerometer-measured physical activity in order to address the question of whether a larger proportion of higher intensity activity is associated with lower adiposity.

10.2.3 Accelerometer and questionnaire measures of physical activity in relation to body size and composition

Prior epidemiologic studies on the associations between physical activity, adiposity, and breast cancer risk have been limited by measurement error for self-reported physical activity[10]. In Chapter 6, I first aimed to characterise and compare the associations between physical activity, measured by questionnaire versus accelerom-

eter, and adiposity in the largest study to date. Physical activity was measured using accelerometers on a subset of the cohort 3-5 years after the recruitment period. The analyses in Chapter 6 expanded the analysis from Chapter 5 which established that greater self-reported physical activity was associated with lower adiposity across all indices of body size and composition by demonstrating that the inverse associations were nearly twofold larger when physical activity was measured by accelerometer compared to questionnaire. Furthermore, there was no plateau in the association between physical activity and adiposity, in the upper ranges of physical activity, when physical activity was measured by accelerometer compared to questionnaire. The results provide strong evidence of a greater inverse association between accelerometer-measured physical activity and adiposity compared to self-reported physical activity and adiposity, in line with previous smaller studies[162,165,166].

The analyses in this thesis extend prior findings by demonstrating that the bias related to self-reported physical activity may be dependent on certain sociodemographic characteristics. For example, although the overall correlation between questionnaire and accelerometer measures of physical activity was low, the correlations were even lower in certain groups such as in participants with higher BMI and in older participants, which suggests a greater magnitude of measurement error in self-reported physical activity among these groups. Additionally, there was heterogeneity by age in the association between body fat percentage and self-reported physical activity but not in the association with accelerometer-measured physical activity, suggesting differential reporting error.

10.2.4 Gene by environment interactions in adiposity

Chapter 7 was a cross-sectional analysis investigating whether the effect of genetic susceptibility, as measured by a genetic risk score of 70 BMI-associated single nucleotide polymorphisms, on objectively measured BMI was modified by physical activity as well as leisure screen time behaviours. Within each fifth of genetic risk score, those who were in the most active group had the lowest BMI, followed by those who were moderately active. Those who were least active had the highest BMI at each level of genetic risk score. Being in the low activity group was associated with a greater excess BMI compared to being in the high activity group, when comparing the top to the bottom reference fifth of genetic risk score. The associations between BMI and physical activity, measured by accelerometer compared to self-reported physical activity, provided further evidence that the genetic predisposition to adiposity can largely be overcome by higher levels of physical activity. Simultaneously, another research group found similar results within UK Biobank[196], although they did not look directly at the associations between physical activity, measured by accelerometer versus questionnaire, in relation to BMI, stratified by genetic risk score.

Being in the high leisure screen time group, which represents sedentary behaviour, was associated with a greater excess BMI compared to being in the group with the lowest leisure screen time, even after adjustment for physical activity level. Both physical activity level and sedentary behaviour were effect modifiers of the association between genetic risk score and BMI, which is consistent with a prior study that found an independent interaction of genetic risk score with physical activity and sedentary behaviour[189]. However, measurement error associated

with self-reported estimates of physical activity and sedentary time is an important methodological issue to take into account when interpreting these results[199].

10.2.5 Indices of adiposity and breast cancer risk in postmenopausal women

Adiposity is a well-established risk factor for postmenopausal breast cancer[122,142,240] likely due to increased oestrogen-stimulated carcinogenesis since adipose tissue is the main site of oestrogen synthesis in postmenopausal women[11,201]. However, only a few smaller studies have investigated whether direct measures of fat mass are better indicators of risk than anthropometric measures[207–210]. The results presented in Chapter 8 constitute the largest prospective study to date to assess whether direct measures of fat mass capture adiposity better than indirect measures, such as BMI, in relation to breast cancer risk in postmenopausal women. There was a continuous increase in breast cancer risk with increasing adiposity, for all measures of body size and composition. Although the associations with breast cancer risk were slightly stronger for body and trunk fat mass, the relative risks, per standard deviation, across the measures of body size and composition were similar and explained a similar degree of variation in risk. This suggests that impedance measures of fat were not substantially better indicators of risk compared to anthropometric measures such as BMI and waist circumference.

The literature is inconsistent on whether central adiposity is associated with breast cancer risk independently of overall adiposity. A recent study reported that waist circumference was independently associated with breast cancer risk after adjustment for BMI[241], yet the other studies included in the most recent meta-

analysis showed no association after adjustment for BMI[6]. I initially hypothesized that these apparent variations in findings may be due in part to the inability of BMI to capture overall body fat, and the self-reported nature of the body size measures in most prior studies. Therefore, when examining the role of abdominal markers of adiposity, I adjusted instead for measured body fat mass, although it only showed a marginally greater association with postmenopausal breast cancer risk, among all the body size and composition variables. I did not find evidence for an independent association between central adiposity and postmenopausal breast cancer risk: after accounting for overall adiposity, the associations between anthropometric measures of central adiposity (waist circumference, hip circumference, and waist-to-hip ratio) and breast cancer risk were attenuated and non-significant. However, it is important to note that because body fat percentage and trunk fat percentage by impedance were very highly correlated, the effects of these two indices of body fatness may be indistinguishable.

In terms of mechanisms, prior studies have reported that breast cancer risk is positively associated with oestradiol[242,243] and testosterone[242]. Obese women had more than twofold higher serum free oestradiol concentrations compared to normal weight women and incorporating serum free oestradiol concentrations in the analysis attenuated the association between BMI and breast cancer risk to the null[22].

10.2.6 Physical activity and breast cancer risk

Previous studies of the association between physical activity and breast cancer risk have been based on self-reported measures of physical activity. In light of

the reporting bias intrinsic to studies using self-reported physical activity, demonstrated in Chapter 6, the analyses in Chapter 9 utilised the accelerometer data available in UK Biobank and constitute the first large-scale prospective, observational analysis to incorporate objective measures of physical activity in examining the association between physical activity and breast cancer risk. The analyses in Chapter 9 clarified the association between physical activity and breast cancer risk using objective accelerometer-based measures in tests for trend, which were generated by using the accelerometer-based median within each baseline physical activity category as a surrogate for questionnaire-based activity.

In multivariable models, accelerometer-based physical activity was associated with a similar reduction in breast cancer risk in both premenopausal and postmenopausal women. Studies on the association between physical activity and breast cancer risk have not consistently examined how the role of adiposity as a potential mediator may differ by menopausal status. In a recent meta-analysis[44], only two of the included studies provided multivariable-adjusted relative risk estimates both before and after adjustment for adiposity and both of these studies are limited to postmenopausal women[227,228]. In agreement with the two prior studies, analyses in this thesis demonstrate that adjusting for body fat mass attenuated the inverse association in postmenopausal women. However, adjusting for body fat mass slightly strengthened the inverse association in premenopausal women. Although the protective association between greater physical activity and breast cancer risk was partly mediated by reduced adiposity in postmenopausal women, these findings overall suggest that other non-adiposity related mechanisms underlie the protective association. While the exact mechanisms remain unclear, several pathways such as reduced concentrations of endogenous sex hormones, improved

insulin sensitivity, enhanced immune function, reduced oxidative stress, reduced genomic instability, and reduced chronic low-grade inflammation have been suggested[48].

Results from recent meta-analyses suggest that the protective association between physical activity and breast cancer risk may be stronger in certain subgroups such as women with a lower BMI compared to women who are obese[44,225]. In contrast, after incorporating accelerometer-based physical activity in tests for trend, analyses in this thesis found no heterogeneity by BMI, menopausal status, hormone replacement therapy use in postmenopausal women, number of births, or alcohol intake frequency. The differences in magnitude of risk reduction seen across the BMI groups in prior studies[225] are likely due to reporting bias. Comparison of self-reported and accelerometer-measured physical activity by BMI group demonstrated substantial over-reporting of physical activity in women with higher BMI. However, it is also possible that there were insufficient numbers to test for heterogeneity by BMI, hormone replacement therapy use, numbers of births, and alcohol intake frequency. Although the analyses presented in this thesis are well-powered for cross-sectional analyses and are among the largest prospective studies of risk factors for breast cancer, the cohort is relatively underpowered for subgroup analyses.

10.3 Methodological considerations

There are several methodological considerations for the research presented in this thesis that have not been discussed in depth within the results chapters themselves, detailed here. For a description of the limitations of the cancer registry data, please

see the methods section in Chapter 8.

10.3.1 Measurement error

Information on covariates was largely from responses to the UK Biobank baseline questionnaire. It is likely that the values for some of these exposures such as alcohol intake, smoking status, and hormone replacement therapy use have changed throughout the duration of follow-up, which would cause some measurement error. However, this is unlikely to have a large impact on the relative risks reported in this thesis given that adjustment for these lifestyle variables made little difference to the estimates.

Furthermore, as discussed in detail in Chapter 3, UK Biobank has collected and continues to collect repeat assessments of the exposures on the baseline questionnaire, in representative subsets, in order to address the "regression dilution" effect that can result from measurement error. Repeat or enhanced measures of baseline exposures have been utilised in the chapters containing detailed, prospective analyses: adiposity and breast cancer risk (Chapter 8) and physical activity and breast cancer risk (Chapter 9).

For the chapter on gene-environment interactions, the measurement error associated with the exposures (physical activity and screen time behaviour) in the main analysis is likely to attenuate the estimate of the true effect size[199]. Therefore, I also used accelerometer-measured physical activity, in addition to the self-reported data, in order to more clearly characterise the association between physical activity and adiposity by genetic risk score.

BMI, as the outcome, was accurately measured by trained staff, but is an

imperfect measure of adiposity. However, I did not use measures of direct body fat instead of BMI in the analyses since the SNPs that comprised the genetic risk score were specific for BMI.

The definition of menopausal status throughout follow-up is likely to be affected by some degree of measurement error since repeat assessment of this information was not available on the entire cohort. Additionally, menopause can be a gradual transition and menopausal status can be masked by hysterectomy or HRT use before menstrual periods naturally stop[84]. Results of sensitivity analyses excluding women from the premenopausal analysis after they reached age 50 did not affect the main findings.

The cancer outcome data came from cancer registries and are dependent on routine coding systems that are likely affected by some degree of misclassification. However, this misclassification is likely to be at random, thus introducing random error instead of bias.

10.3.2 Confounding

Observational epidemiologic studies are potentially limited by confounding. A confounder is a variable that is associated with the exposure and causally associated with the outcome of interest that can bias the observed association between exposure and outcome. The analyses presented in this thesis have been stratified and adjusted for a large number of sociodemographic factors and established risk factors for breast cancer for which information was available. One established risk factor for which information was not available is breastfeeding, which is protective[78], particularly against hormone receptor-negative breast cancers[244]. However,

breastfeeding is unlikely to be a major confounder, given that the analyses have been stratified by socioeconomic status and adjusted for parity. If confounders are not well-measured, adjustment is incomplete and can result in residual confounding.

10.3.3 Reverse causality

Reverse causality occurs when preclinical disease causes changes in behavioural or physiological characteristics, potentially giving rise to spurious associations. Although this has been discussed in the context of cross-sectional studies in prior chapters, reverse causality can also be a concern in prospective analyses. For example, the association between physical activity and breast cancer risk could plausibly be influenced by reverse causality if participants reduce their level of physical activity due to feeling ill prior to diagnosis of breast cancer. However, sensitivity analyses excluding the first two years of follow-up and excluding those who reported poor health or a long-term illness or disability did not differ substantially from the main results, suggesting reverse causality was not an issue in these analyses.

10.3.4 Generalisability and representativeness

The UK Biobank cohort is not representative of the sampling population. The participation rate in women was 6.4%[114]. Overall, a comparison of the sociodemographic and health-related characteristics of study participants to those of the general population provided evidence of a "healthy volunteer" selection bias: participants tended to live in less socioeconomically deprived areas, were less likely to

be obese, were less likely to drink alcohol daily, were less likely to smoke, and had lower rates of total cancer incidence and all-cause mortality[114]. However, results of associations between exposures and health outcomes may be generalisable to other populations and would not require the study population to be representative if the biological basis of the exposure-disease relationship is shared.

However, there are some instances where the issues around representativeness and generalisability must be considered more specifically. For example, participants in the UK Biobank were predominately of white ethnicity[114] and participants of non-European ancestry were specifically excluded from the analyses using the genetic data, given that the genome-wide significant loci used to create the genetic risk score in Chapter 7 were based on analyses restricted to individuals of European descent[170]. Therefore, the results from this thesis on gene-environment interactions in adiposity may not be generalisable to populations of differing ethnicity.

Since information was not available to run analyses according to hormone receptor status, it should also be noted that the results seen in this thesis may be largely reflective of hormone receptor positive breast cancers[222], and therefore may not be generalisable to other tumour subtypes.

10.4 Public health implications

Potentially modifiable risk factors for breast cancer likely underlie much of the differences in incidence rates seen globally, as evidenced by migration studies in which breast cancer incidence rates increased in descendants of women who moved from low to high incidence areas to levels comparable to the high incidence areas[245].

Moreover, a recent study of population attributable risk of various breast cancer risk factors in the Nurses' Health Study suggested that more than a third of postmenopausal breast cancers are preventable through risk factors that are modifiable at menopause[121]. Therefore, the identification and characterisation of modifiable risk factors for breast cancer is of utmost importance for breast cancer prevention and risk reduction.

After accounting for height and other confounders, being in the top compared to bottom quartile of body fat mass was associated with a 70% increased risk of breast cancer among postmenopausal women who were not current users of HRT. This finding highlights the maintenance of healthy body size and composition as the most important modifiable risk factor for postmenopausal breast cancer, in terms of attributable risk in populations with high levels of overweight and obesity.

The benefits of physical activity in relation to adiposity may be greater than previously estimated, as suggested by the nearly twofold greater inverse association between physical activity and adiposity, when measured by accelerometer compared to questionnaire. The approximately linear dose-response relationship between accelerometer-measured physical activity and adiposity suggests a continuously greater benefit with increasing physical activity.

The finding that greater physical activity attenuated the association of genetic risk score with BMI and obesity risk demonstrates that being physically active can largely overcome genetic predisposition to obesity. Individuals with an adverse genetic risk profile for adiposity can benefit considerably from adopting or maintaining a physically active lifestyle. Moreover, the association of genetic risk score with greater BMI was stronger in the setting of high sedentary activity, which raises the population-level issue of whether those with a greater genetic predispo-

sition for adiposity may be increasingly susceptible to obesity as lifestyles become more sedentary. Further public health efforts are needed to combat sedentary lifestyles, which are highly prevalent in Europe[246].

The protective association between physical activity and breast cancer risk in both premenopausal and postmenopausal women emphasises the importance of maintaining an active lifestyle. Furthermore, the accelerometer-based trends provided no evidence that physical activity is associated with a smaller breast cancer risk reduction in the women with higher BMI. Rather, these results demonstrate that all women, regardless of body size, would benefit from being physically active.

10.5 Recommendations for further research

An immediate next step for the research presented in this thesis includes estimating the public health impact of the exposures in the study population. This can be done by calculating the population attributable risks.

In the longer term, the work presented in this thesis should be extended by looking at whether the associations between body size and composition, physical activity, and breast cancer risk are mediated by levels of endogenous hormones and growth factors such as oestrogen, testosterone, and insulin-like growth factor 1. Greater physical activity may be associated with lower concentrations of testosterone in premenopausal women and lower oestradiol as well as higher sex hormone binding globulin concentrations in postmenopausal women, independent of BMI[247]. However, further studies are required with better assays in order to provide more accurate estimates for the associations of endogenous hormones and growth factors in relation to breast cancer risk. More specifically, more evidence is

needed in relation to the possible association between oestradiol and breast cancer risk in premenopausal women[90] and also to clarify how the associations may vary by hormone receptor status. In UK Biobank, biological samples were collected at baseline from all 0.5 million participants (over half of whom are women), making this data a valuable resource among many other large prospective cohort studies.

While follow-up time after collection of accelerometer measurements was too short to study incident breast cancer in the analyses presented in this thesis, future studies should prospectively investigate the association between objective accelerometer-measured physical activity, preferably with repeat measurements, and risk of incident breast cancer. Given that accelerometers are worn continuously every hour of the day, patterns of physical activity and sedentary behaviour should be characterised together in the same joint framework, rather than in isolation. Moreover, accelerometer-measured physical activity has the potential to more accurately characterise the role of frequency, duration, intensity and patterns of physical activity in relation to adiposity and breast cancer risk. The accelerometer-measured physical activity variable currently available in UK Biobank cannot be directly compared to MET hours of self-reported physical activity and is difficult to interpret as a basis for public health messages. However, as mentioned in Chapter 6, physical activity phenotypes have been developed. Once these variables are made publically available in UK Biobank, research using these metrics will facilitate the translation of study results into public health messages.

The analyses in this thesis have utilised impedance and DXA measures, but to gain a deeper understanding of the associations between adiposity and health outcomes, future research on body size and composition measures should utilise magnetic resonance imaging, which provides more accurate estimates of body fat

content and distribution. This method allows for the distinct quantification of visceral adipose tissue and abdominal subcutaneous adipose tissue[248]. Magnetic resonance-based body composition measurement and analysis are currently ongoing in UK Biobank[249].

10.6 Conclusion

In conclusion, the results presented in this thesis suggest that physical activity is important not only for maintaining body size and composition but also for reducing breast cancer risk. Physical activity may provide more public health benefit than previously thought, as evidenced by the greater protective associations seen with objective measures compared to self-reported measures of physical activity. Physical activity and inactivity modify the effect of genetic predisposition to adiposity. The association of objectively-measured physical activity with adiposity provides further evidence that genetic predisposition to overweight and obesity can be largely overcome with physical activity.

Future research should focus on exploring the mechanisms underlying the protective association of physical activity and breast cancer risk, beyond the role of adiposity in postmenopausal breast cancer. From the public health perspective, further efforts are needed to raise awareness of the importance of maintaining a healthy body size in reducing postmenopausal breast cancer risk as well as the benefits of physical activity on adiposity and breast cancer risk in both premenopausal and postmenopausal women.

Appendix A

QUESTIONS FROM UK BIOBANK TOUCHSCREEN QUESTIONNAIRE RELEVANT TO THIS THESIS

Appendix A: Questions from the UK Biobank touchscreen (baseline) questionnaire relevant to this thesis

Question stem	Possible responses
Which of the following describes your current situation? (You can select more than one answer)	• In paid employment or self-employed • Retired • Looking after home and/or family • Unable to work because of sickness or disability • Unemployed • Doing unpaid or voluntary work • Full or part-time student • None of the above • Prefer not to answer
Does your work involve walking or standing for most of the time?	"Never/rarely" OR "Sometimes" OR "Usually" OR "Always" OR "Do not know" OR "Prefer not to answer"
Does your work involve heavy manual or physical work?	"Never/rarely" OR "Sometimes" OR "Usually" OR "Always" OR "Do not know" OR "Prefer not to answer"
Which of the following qualifications do you have? (You can select more than one)	• College or university degree • A levels/AS levels or equivalent • O levels/GCSEs or equivalent • CSEs or equivalent • NVQ or HND or HNC or equivalent • Other professional qualifications eg: nursing, teaching • None of the above

Intro the physical activity questions: Next are some questions about your day to day activities. We know it may be difficult to answer some of these questions exactly but an approximate answer is better than none. The first few questions ask about the time you have spent being physically active in the last 4 weeks. To answer these questions, think about all the activities that you do at work, as part of your house work, to get from place to place and in your leisure time. Please touch 'next' to continue.

Question stem	Possible responses
In a typical WEEK, on how many days did you walk for at least 10 minutes at a time? (include walking that you do at work, travelling to and from work, and for sport or leisure)	Enter number OR "Do not know" OR "Unable to walk" or "Prefer not to answer"
How many minutes did you usually spend walking on a typical DAY?	Enter number OR "Do not know" OR "Prefer not to answer" (Note: this question was skipped if the participant answered "unable to walk" or "do not know" or 0 on the previous question)
In a typical WEEK, on how many days did you do 10 minutes or more of moderate physical activities like carrying light loads, cycling at normal pace? (do not include walking)	Enter number OR "Do not know" OR "Prefer not to answer"
How many minutes did you usually spend doing moderate activities on a typical DAY?	Enter number OR "Do not know" OR "Prefer not to answer" (Note: this question was skipped if the participant answered "Do not know" or 0 or "prefer not to answer" on the previous question)
In a typical WEEK, how many days did you do 10 minutes or more of vigorous physical activity? (these are activities that make you sweat or breathe hard such as fast cycling, aerobics, heavy lifting)	Enter number OR "Do not know" OR "Prefer not to answer"
How many minutes did you usually spend doing vigorous activities on a typical DAY?	Enter number OR "Do not know" OR "Prefer not to answer" (Note: this question was presented only if the participant gave a value greater than 0 on the previous question)

Question stem	Possible responses
In a typical DAY, how many hours do you spend watching TV? Include time spent watching videos & DVDs. (Put 0 if you do not spend any time doing it)	Enter number OR “Less than an hour a day” OR “Do not know” OR “Prefer not to answer”
In a typical DAY, how many hours do you spend using the computer (do not include using a computer at work; put 0 if you do not spend any time doing it)	Enter number OR “Less than an hour a day” OR “Do not know” OR “Prefer not to answer”
Do you smoke tobacco now?	“Yes, on most or all days” OR “Only occasionally” OR “No” or “Prefer not to answer”
How old were you when you last smoked cigarettes on most days?	Enter number OR “Do not know” OR “Prefer not to answer”

Question stem	Possible responses
About how often do you drink alcohol? (Note: Participants who answered “special occasions only” or “one to three times a month” were subsequently asked about alcohol intake amount in an average MONTH. Participants who answered “once or twice a week”, “three or four times a week”, or “daily or almost daily” were subsequently asked about alcohol intake amount in an average WEEK.)	Select one from • Daily or almost daily • Three or four times a week • Once or twice a week • One to three times a month • Special occasions only • Never • Prefer not to answer
In an average MONTH (or WEEK, where appropriate), how many glasses of RED wine would you drink? (there are six glasses in an average bottle)	Enter number OR “Do not know” OR “Prefer not to answer”
In an average MONTH (or WEEK, where appropriate), how many glasses of WHITE wine or champagne would you drink? (there are six glasses in an average bottle)	Enter number OR “Do not know” OR “Prefer not to answer”
In an average MONTH (or WEEK, where appropriate), how many pints of beer or cider would you drink? (Include bitter, lager, stout, ale, Guinness)	Enter number OR “Do not know” OR “Prefer not to answer”
In an average MONTH (or WEEK, where appropriate), how many measures of spirits or liqueurs would you drink? (there are 25 standard measures in a normal sized bottle; spirits include drinks such as whisky, gin, rum, vodka, brandy)	Enter number OR “Do not know” OR “Prefer not to answer”
In an average MONTH (or WEEK, where appropriate), how many glasses of fortified wine would you drink? (there are 12 glasses in an average bottle). (Fortified wines include drinks such as sherry, port, vermouth)	Enter number OR “Do not know” OR “Prefer not to answer”
In an average MONTH (or WEEK, where appropriate), how many glasses of other alcoholic drinks (such as alcopops) would you drink?	Enter number OR “Do not know” OR “Prefer not to answer”

Question stem	Possible responses
What is your ethnic group?	"White" OR "Mixed" OR "Asian or Asian British" OR "Black or Black British" OR "Chinese" OR "Other ethnic group" "Do not know" or "Prefer not to answer"
When you were 10 years old, compared to average would you describe yourself as:	"Thinner" OR "Plumper" OR "About average" OR "Do not know" OR "Prefer not to answer"
Has/did your mother ever suffer from? (You can select more than one answer)	"Parkinson's disease", "Severe depression", "Lung cancer", "Bowel cancer", **"Breast cancer**", "None of the above", "Do not know", "Prefer not to answer"
Have any of your brothers or sisters suffered from any of the following illnesses? (You can select more than one)	"Parkinson's disease", "Severe depression", "Lung cancer", "Bowel cancer", **"Breast cancer**", "None of the above", "Do not know", "Prefer not to answer"
In general how would you rate your overall health?	"Excellent" OR "Good" OR "Fair" OR "Poor" OR "Do not know" OR "Prefer not to answer"
Do you have any long-standing illness, disability or infirmity?	"Yes" OR "No" OR "Do not know" OR "Prefer not to answer"

Question stem	Possible responses
Have you ever been for breast cancer screening (a mammogram)?	"Yes" OR "No" OR "Do not know" OR "Prefer not to answer"
How old were you when your periods started?	Enter number OR "Do not know" OR "Prefer not to answer"
Have you had your menopause (periods stopped)?	"Yes" OR "No" OR "Not sure – had a hysterectomy" OR "Not sure – other reason" OR "Prefer not to answer"
How old were you when your periods stopped?	Enter number OR "Do not know" OR "Prefer not to answer"
How many children have you given birth to? (please include live births only)	Enter number OR "Prefer not to answer"
How old were you when you had your child?	Enter number OR "Do not remember" OR "Prefer not to answer"
How many stillbirths? (enter 0 if none)	Enter number OR "Do not remember" OR "Prefer not to answer" (Note: this was only asked if the participant indicated she had ever had any "stillbirths, spontaneous miscarriages or terminations"
Have you ever taken the contraceptive pill? (include the 'mini- pill')	"Yes" OR "No" OR "Do not know" OR "Prefer not to answer"
Have you ever used hormone replacement therapy (HRT)?	"Yes" OR "No" OR "Do not know" OR "Prefer not to answer"
How old were you when you last used HRT?'	Enter number OR "Do not know" OR "Still taking HRT" OR "Prefer not to answer"
Have you had BOTH ovaries removed?	"Yes" OR "No" OR "Do not know" OR "Prefer not to answer"

Appendix B

INTERNATIONAL PHYSICAL ACTIVITY QUESTIONNAIRE SCORING GUIDELINES

7.4 Truncation of Data Rules

This rule attempts to normalize the distribution of levels of activity which are usually skewed in national or large population data sets.

In IPAQ short - it is recommended that all Walking, Moderate and Vigorous time variables exceeding ' 3 hours' or '180 minutes' are truncated (that is re-coded) to be equal to '180 minutes' in a new variable. This rule permits a maximum of 21 hours of activity in a week to be reported for each category (3 hours * 7 days).

In IPAQ long – the truncation process is more complicated, but to be consistent with the approach for IPAQ short requires that the variables total Walking, total Moderate-intensity and total Vigorous-intensity activity are calculated and then, for each of these summed behaviours, the total value should be truncated to 3 hours (180 minutes).

When analysing the data as categorical variable or presenting median and interquartile ranges of the MET-minute scores, the application of the truncation rule will not affect the results. This rule does have the important effect of preventing misclassification in the 'high' category. For example, an individual who reports walking for 10 minutes on 6 days and 12 hours of moderate activity on one day could be coded as 'high' because this pattern meets the '7 day" and "3000 MET-min" criteria for 'high'. However, this uncommon pattern of activity is unlikely to yield the health benefits that the 'high' category is intended to represent.

Although using median is recommended due to the skewed distribution of scores, if IPAQ data are analysed and presented as a continuous variable using mean values, the application of the truncation rule will produce slightly lower mean values than would otherwise be obtained.

7.5 Calculating MET-minute/week Scores

Data processing rules 7.2, 7.3, and 7.4 deals first with excluding outlier data, then secondly, with recoding minimum values and then finally dealing with high values. These rules will ensure that highly active people remain classified as 'high', while decreasing the chances that less active individuals are misclassified and coded as 'high'.

Using the resulting variables, convert time and days to MET-minute/week scores [see above Sections 5.2 and 6.2; METS x days x daily time].

7.6 Calculating Total Days for Presenting Categorical Data on Moderate and High Levels

Presenting IPAQ data using categorical variables requires the total number of 'days' on which all physical activity was undertaken to be assessed. This is difficult because frequency in 'days' is asked separately for walking, moderate-intensity and vigorous-intensity activities, thus allowing the total number of 'days' to range from a minimum

of 0 to a maximum of 21'days' per week in IPAQ short and higher in IPAQ long. The IPAQ instrument does not record if different types of activity are undertaken on the same day.

In calculating 'moderately active', the primary requirement is to identify those individuals who undertake activity on at least '5 days'/week [see Sections 4.2 and 5.3]. Individuals who meet this criterion should be coded in a new variable called "*at least five days"* and this variable should be used to identify those meeting criterion b) at least 30 minutes of moderate-intensity activity and/or walking; and those meeting criterion c) any combination of walking, moderate-intensity or vigorous-intensity activities achieving a minimum of 600 MET-minutes/week.

Below are two examples showing this coding in practice:

i) an individual who reports '2 days of moderate-intensity' and '3 days of walking' should be coded as a value indicating "*at least five days"*;
ii) an individual reporting '2 days of vigorous-intensity', '2 days of moderate-intensity' and '2 days of walking should be coded as a value to indicate "*at least five days"* [even though the actual total is 6].

The original frequency of 'days' for each type of activity should remain in the data file for use in the other calculations.

The same approach as described above is used to calculate total days for computing the 'high' category. The primary requirement according to the stated criteria is to identify those individuals who undertake a combination of walking, moderate-intensity and or vigorous-intensity activity on at least 7 days/week [See section 4.2]. Individuals who meet this criterion should be coded as a value in a new variable to reflect *"at least 7 days*".

Below are two examples showing this coding in practice:

i) an individual who reports '4 days of moderate-intensity' and '3 days of walking' should be coded as the new variable *"at least 7 days*".
ii) an individual reporting '3 days of vigorous-intensity', '3 days moderate-intensity' and '3 days walking' should be coded as *"at least 7 days*" [even though the total adds to 9] .

8. Summary algorithms

The algorithms in Appendix 1 and Appendix 2 to this document show how these rules work in an analysis plan, to develop the categories 1 [Low], 2 [Moderate], and 3 [High] levels of activity.

IPAQ Research Committee
November 2005

APPENDIX 1

At A Glance
IPAQ Scoring Protocol (Short Forms)

Continuous Score

Expressed as MET-min per week: MET level x minutes of activity/day x days per week

Sample Calculation

MET levels	**MET-minutes/week for 30 min/day, 5 days**
Walking = 3.3 METs	3.3*30*5 = 495 MET-minutes/week
Moderate Intensity = 4.0 METs	4.0*30*5 = 600 MET-minutes/week
Vigorous Intensity = 8.0 METs	8.0*30*5 = 1,200 MET-minutes/week
	TOTAL = 2,295 MET-minutes/week

Total MET-minutes/week = Walk (METs*min*days) + Mod (METs*min*days) + Vig (METs*min*days)

Categorical Score- three levels of physical activity are proposed

1. <u>Low</u>

- No activity is reported **OR**
- Some activity is reported but not enough to meet Categories 2 or 3.

2. <u>Moderate</u>

Either of the following 3 criteria
- 3 or more days of vigorous activity of at least 20 minutes per day **OR**
- 5 or more days of moderate-intensity activity and/or walking of at least 30 minutes per day **OR**
- 5 or more days of any combination of walking, moderate-intensity or vigorous-intensity activities achieving a minimum of at least 600 MET-minutes/week.

3. <u>High</u>

Any one of the following 2 criteria
- Vigorous-intensity activity on at least 3 days and accumulating at least 1500 MET-minutes/week **OR**
- 7 or more days of any combination of walking, moderate- or vigorous-intensity activities accumulating at least 3000 MET-minutes/week

Please review the full document "Guidelines for the data processing and analysis of the International Physical Activity Questionnaire" for more detailed description of IPAQ analysis and recommendations for data cleaning and processing [www.ipaq.ki.se].

Appendix C

PAPERS PUBLISHED FROM DPHIL WORK

C.1 Physical activity in relation to body size and composition in women in UK Biobank (Annals of Epidemiology, 2015)

I am the first author of this paper, which is based on work in Chapter 5. I conducted the statistical analyses, drafted the paper, and revised it to address reviewers' comments, with input from the co-authors.

Annals of Epidemiology 25 (2015) 406–413

Contents lists available at ScienceDirect

Annals of Epidemiology

journal homepage: www.annalsofepidemiology.org

Original article

Physical activity in relation to body size and composition in women in UK Biobank

Wenji Guo BS *, Kathryn E. Bradbury PhD, Gillian K. Reeves PhD, Timothy J. Key PhD

Cancer Epidemiology Unit, Nuffield Department of Population Health, University of Oxford, Oxford, UK

ARTICLE INFO

Article history:
Received 8 August 2014
Accepted 25 January 2015
Available online 4 February 2015

Keywords:
Adiposity
Body fat percentage
Body mass index
Cross-sectional studies
Intensity
Menopause
Physical activity

ABSTRACT

Purpose: Physical activity is thought to protect against obesity, but the evidence is limited and few large studies of this topic have direct, objective measures of several different anthropometric variables. We examined the association of self-reported physical activity with measures of total and central adiposity.
Methods: Our cross-sectional study included 38,201 premenopausal and 94,592 postmenopausal healthy women aged 40 to 70 years in UK Biobank, recruited from 2006 to 2010.
Results: Means for total and vigorous physical activity were 31.3 (SD, 30.8) and 9.7 (SD, 14.8) metabolic equivalent h/wk, respectively, for premenopausal and 34.4 (SD, 33.1) and 8.9 (SD, 15.6) metabolic equivalent h/wk, respectively, for postmenopausal women. Multiple linear regression models showed that in both premenopausal and postmenopausal women, body mass index, body fat mass and percentage, trunk fat mass and percentage, waist and hip circumference, and waist-to-hip ratio were lower in women with higher physical activity ($P < .0001$). Within each category of total physical activity, a larger proportion of vigorous physical activity was associated with lower adiposity ($P < .0001$).
Conclusions: These findings support the roles of a physically active lifestyle and vigorous exercise in maintaining healthy body size and composition. Higher exercise intensity may be associated with lower adiposity, beyond the influence of exercise frequency and duration.
© 2015 Elsevier Inc. All rights reserved.

Introduction

Obesity is of major, general public health concern because of its association with increased risk of a variety of diseases including diabetes, heart disease, hypertension, stroke, and certain cancers such as breast and endometrial [1]. Although physical activity is widely believed to be an important determinant of body size and composition, the full nature of the relationship is still unclear [2,3]. Although observational studies support an association between exercise and weight-loss maintenance, randomized controlled trials have produced inconsistent results [4]. Questions also remain over the role of vigorous compared with lower intensity physical activity in reducing adiposity and preventing weight gain, and over the relationship between physical activity and different measures of overall versus central adiposity [5,6]. Because risks for some obesity-related diseases such as breast cancer and cardiovascular disease differ by menopausal status, it is of particular interest to examine the associations between physical activity and adiposity separately according to menopausal status in women [7,8]. Some previous large studies have been limited by low numbers of women reporting vigorous physical activity [9].

A number of anthropometric variables have been reported and used in public health recommendations. Because an increase in particular indices of body size may have an effect on specific health outcomes, it is important to assess which indices are likely to be modified by behaviors such as exercise. We report here a cross-sectional analysis of associations between reported levels of overall and of vigorous physical activity and several objective measures of body size and composition in more than 132,000 premenopausal and postmenopausal women. In particular, we assess the hypothesis that physical activity is related to body size and composition and that associations may vary by particular measures of adiposity. This research has been conducted using the UK Biobank resource, a nationwide study of approximately 500,000 UK adults.

Methods

Data source

Data were obtained by application to UK Biobank (reference number 3248, approved in August 2013). Details of the rationale,

* Corresponding author. Cancer Epidemiology Unit, Nuffield Department of Population Health, University of Oxford, Richard Doll Building, Roosevelt Drive, Oxford OX3 7LF, UK. Tel.: +44 (0) 1865 289 600; fax: +44 (0) 1865 289 610.
E-mail address: wenji.guo@ceu.ox.ac.uk (W. Guo).

http://dx.doi.org/10.1016/j.annepidem.2015.01.015
1047-2797/© 2015 Elsevier Inc. All rights reserved.

design, and survey methods for UK Biobank have been described elsewhere [10], and information on data available and access procedures are given on the study Web site (http://www.ukbiobank.ac.uk/). UK Biobank has approval from the North West Multi-centre Research Ethics Committee, the National Information Governance Board for Health and Social Care in England and Wales, and the Community Health Index Advisory Group in Scotland. All participants provided written informed consent.

Study participants

The complete UK Biobank data set includes 502, 656 UK adults (229,182 men and 273,474 women) between 40 to 70 years of age at recruitment during 2006 to 2010. Because participation was voluntary and the data collection process rigorous, UK Biobank aimed to enhance generalizability rather than to be representative of the UK population. Participant recruitment was conducted via centrally coordinated identification and invitation from population-based National Health Service patient registers of individuals living within a reasonable distance of an assessment center. Participants completed a touchscreen questionnaire during the baseline assessment center visit that included questions on sociodemographics, lifestyle, health and medical history, and sex-specific factors. Menopausal status provided by UK Biobank was self-reported; for these analyses, women without self-reported menopausal status who were older than 53 years at recruitment and/or had had both ovaries removed were categorized as postmenopausal ($n = 29,777$), and women younger than 45 years who had not undergone a bilateral oophorectomy were categorized as premenopausal ($n = 1053$).

Women were excluded from the analyses if they reported at recruitment a long-standing illness, disability, infirmity, or had unknown status for this variable ($n = 87,989$), had unknown menopausal status ($n = 15,517$), were pregnant ($n = 121$), or had missing data on any of the physical activity variables used in our analyses ($n = 70,343$). Participants were also excluded if the sum of all walking, moderate, and vigorous time variables was greater than 6720 min/wk as detailed in the International Physical Activity Questionnaire (IPAQ) scoring guidelines ($n = 2184$), which can be accessed at https://sites.google.com/site/theipaq/scoring-protocol. Included in the analyses were 38,201 premenopausal and 94,592 postmenopausal women.

Physical activity

Physical activity questions asked on the baseline questionnaire captured the frequency and duration of three different levels of activity (walking, moderate, and vigorous). Participants were asked how many days per typical week they spent engaged in each category of activity. For each category in which an answer of one or more days was given, the participant was then asked the number of minutes spent on the activity during a typical day. Questions were adapted from the IPAQ short form, a validated survey instrument [11]. Questions are listed in Supplementary Table 1 and are available to view on the UK Biobank Web site (http://biobank.ctsu.ox.ac.uk/crystal/). We used metabolic equivalents (METs) to quantify physical activity; 1 MET is expended by sitting quietly for 1 hour, and the MET value reflects the ratio of energy expended per kilogram of body weight per hour to that expended when sitting quietly [12]. The number of minutes per day engaged in each level of activity was multiplied by the respective MET score for that level of activity. MET minutes per day were converted to MET hours per week. The total amount of METs was calculated by summing total METs from each activity level. In line with IPAQ guidelines [11], physical activity of less than 10 min/d for walking, moderate, or vigorous activity was recoded to 0.

Each category was assigned the following MET hour values: 3.3 for walking, 4.0 for moderate physical activity, and 8.0 for vigorous physical activity [11]. Because individuals are estimated to expend 1 MET/h sitting quietly, we report physical activity as excess METs [13]. Excess METs were calculated by subtracting one from the MET value multiplier for each category, giving values of 2.3 for walking, 3.0 for moderate physical activity, and 7.0 for vigorous physical activity.

To assess the importance of vigorous physical activity relative to total physical activity, vigorous physical activity was divided by total physical activity to generate a ratio of vigorous-to-total physical activity. The median ratio of vigorous-to-total physical activity in participants who engaged in any vigorous physical activity was 33%. For analysis, participants were grouped into three categories of vigorous-to-total physical activity: zero percent vigorous physical activity, less than 33% vigorous physical activity, and more than 33% vigorous physical activity.

Anthropometry and body composition

At the UK Biobank baseline interview, trained staff used standard procedures to measure and record standing height using the Seca 202 device (Seca, Hamburg, Germany). The Wessex non-stretchable sprung tape (Wessex, United Kingdom) measure was used to measure waist and hip circumference. The Tanita BC-418MA body composition analyzer (Tanita, Tokyo, Japan) was used to assess bioimpedance measures including body fat percentage, whole body fat mass, trunk fat percentage, and trunk fat mass. UK Biobank derived body mass index (BMI) by dividing weight (kilogram) by the square of standing height (square meter). We derived waist-to-hip ratio (WHR) by dividing waist circumference by hip circumference.

Statistical analyses

Multiple linear regression models were used to examine the association between physical activity and the various anthropometric and body composition measures. Amount of physical activity was categorized into nine groups to obtain an approximately even distribution of participants (0, 1–4.9, 5–9.9, 10–14.9, 15–24.9, 25–34.9, 35–49.9, 50–74.9, and ≥75 excess MET-h/wk) and the figures show the mean within each category. Those who did no physical activity were placed in a separate category even though it contained fewer individuals. For each anthropometric outcome variable, a test for difference of nine means was performed. Participant characteristics are reported for three broad levels of physical activity: low (<10 excess MET-h/wk), moderate (10–49.9 excess MET-h/wk), and high (≥50 excess MET-h/wk). The relationship between total physical activity and anthropometric variables was examined separately for those for whom all physical activity was nonvigorous, less than 33% of total physical activity was vigorous, and more than 33% of total physical activity was vigorous. Ratios of vigorous-to-total physical activity excess MET-hours per wk were derived. To assess whether the association between total physical activity and the anthropometric outcome variable varies significantly depending on proportion of vigorous-to-total physical activity, we performed likelihood ratio tests comparing models which estimated the overall relationship of total physical activity with the corresponding models in which the relationship of total physical activity was allowed to vary according to the ratio of vigorous-to-total physical activity. All analyses were performed separately for premenopausal and postmenopausal women. Analyses were conducted using STATA, version 13.0 (State Corp LP, College Station, TX).

Covariates included 5-year age at recruitment categories, educational qualifications, parity, smoking status, employment status, alcohol intake, and quintiles of Townsend deprivation index

Table 1
Summary characteristics for the 38,201 premenopausal women participants of UK Biobank by PA level

Characteristics	Low PA (<10 MET-h/wk)	Moderate PA (10–49.9 MET-h/wk)	High PA (≥50 MET-h/wk)	All premenopausal women
Number of participants	9142	21,730	7329	38,201
Age, y: mean (SD)	45.6 (3.5)	45.5 (3.4)	45.6 (3.4)	45.5 (3.4)
Total excess MET-h/wk: mean (SD)	5.1 (2.9)	25.3 (10.7)	91.4 (45.9)	31.3 (30.8)
Vigorous excess MET-h/wk: mean (SD)	<1 (1.4)	8.4 (8.1)	25.0 (24.7)	9.7 (14.8)
Height, cm: mean (SD)	164.2 (6.3)	164.5 (6.2)	164.3 (6.2)	164.4 (6.2)
Weight, kg: mean (SD)	72.4 (15.2)	69.1 (12.8)	68.0 (12.2)	69.7 (13.4)
BMI, kg/m^2: mean (SD)	26.8 (5.4)	25.5 (4.5)	25.2 (4.3)	25.8 (4.8)
<18.50, %	<1	<1	<1	<1
18.5–24.99, %	42.4	52.8	55.8	50.7
25.00–29.99, %	33.9	31.5	30.9	32.0
≥30.00, %	23.0	14.9	12.4	16.3
Body fat percentage: mean (SD)	35.8 (7.0)	33.5 (6.8)	32.4 (6.9)	33.8 (7.0)
Body fat mass, kg: mean (SD)	26.8 (10.7)	23.9 (9.2)	22.7 (8.8)	24.4 (9.6)
Trunk fat percentage: mean (SD)	33.4 (7.9)	31.0 (7.8)	29.6 (8.0)	31.3 (7.9)
Trunk fat mass, kg: mean (SD)	13.6 (5.6)	12.0 (5.0)	11.3 (4.8)	12.3 (5.2)
Waist circumference, cm: mean (SD)	83.3 (12.4)	80.0 (10.8)	79.0 (10.6)	80.6 (11.3)
Hip circumference, cm: mean (SD)	103.6 (10.6)	101.1 (9.1)	100.0 (8.8)	101.5 (9.5)
WHR: mean (SD)	0.80 (0.07)	0.79 (0.06)	0.79 (0.06)	0.79 (0.06)
Qualifications* (%)				
College or university degree	44.1	48.7	39.1	45.7
National examination at ages 17–18† y	15.8	15.6	15.2	15.6
National examination at age 16‡ y	31.5	27.9	34.3	30.0
Vocational qualifications	5.8	5.8	7.4	6.1
In paid employment or self-employed	91.1	88.3	86.3	88.6
Standing or walking job§ of all participants	15.0	20.5	45.7	24.0
Manual job‖ of all participants	2.4	4.6	19.4	7.0
Alcohol consumption (%)				
Never or rarely	18.3	15.1	18.8	16.6
Weekly	44.4	43.5	44.4	43.9
More than weekly	37.3	41.3	36.9	39.5
Smoking status (%)				
Never	66.4	65.5	62.9	65.2
Previous	23.2	25.9	26.9	25.5
Current	10.3	8.4	10.0	9.2
Parity: mean (SD)	1.5 (1.2)	1.5 (1.2)	1.7 (1.2)	1.6 (1.2)
Nulliparous, %	26.8	27.1	24.0	26.4
One or two, %	56.6	54.9	54.1	55.2
Three or more, %	16.6	18.0	21.8	18.4

PA = physical activity.
* Numbers do not add up to 100% due to missing data.
† Referred to as A levels, AS levels, or equivalent in UK Biobank.
‡ Referred to as O levels, GCSEs, CSEs, or equivalent in UK Biobank.
§ Participants who reported their work "usually" or "always" involved walking or standing for most of the time.
‖ Participants who reported their work "usually" or "always" involved heavy manual or physical work for most of the time.

[14]. Educational qualifications were classified as vocational qualifications (NVQ, HND, HNC, equivalent, or other professional qualifications), national exams at the age of 16 years (O levels, GCSEs, CSEs, or equivalent), optional national exams at ages 17 to 18 years (A levels, AS levels, or equivalent), and college or university degree. Smoking status was categorized as never, previous, and current. Employment status was classified as paid or self-employment, retired, looking after home and/or family, unemployed, doing unpaid or voluntary work, unable to work due to sickness or disability, and student. Alcohol intake was categorized as never, special occasions only, 1 to 3 times a month, 1 to 2 times a week, 3 to 4 times a week, and daily or almost daily. For postmenopausal women, all analyses were additionally adjusted for hormone replacement therapy use (never, previous, and current). Missing data were grouped in a separate unknown category for each covariate. There were less than 1% missing data for all covariates, with the exception of educational qualifications in postmenopausal women (17%).

Results

Descriptive statistics for anthropometric, physical activity and sociodemographic variables are listed in Table 1 for premenopausal women and Table 2 for postmenopausal women. For all women, the distribution of physical activity was right-skewed. For premenopausal women, the median (interquartile range) physical activity was 22.1 (10.5–41.6), and the mean (SD) was 31.3 (30.8) excess MET-h/wk. For postmenopausal women, the median (interquartile range) was 23.9 (11.3–46.9), and the mean (SD) was 34.4 (33.1) excess MET-h/wk. Although the mean total amount of physical activity was higher in postmenopausal than in premenopausal women, postmenopausal women performed less vigorous physical activity than the premenopausal women (8.9 [15.6] compared with 9.7 [14.8] excess MET-h/wk). Postmenopausal women had higher mean values than premenopausal women for all body composition variables, except for weight, which was nearly identical (on average, premenopausal women were 2 cm taller).

Women who had a high level of physical activity were the least likely to possess a college or university degree, whereas women with a moderate level of physical activity were most likely to hold a college or university degree (Tables 1 and 2). Those with a high level of physical activity were more likely to have manual and/or standing jobs than those with a low or moderate level of physical activity. Supplementary Table 2 compares characteristics of manual and nonmanual workers.

Table 2
Summary characteristics for the 94,592 postmenopausal women participants of UK Biobank by PA level

Characteristics	Low PA (<10 MET-h/wk)	Moderate PA (10–49.9 MET-h/wk)	High PA (≥50 MET-h/wk)	All postmenopausal women
Number of participants	20,756	52,053	21,783	94,592
Age, y: mean (SD)	59.1 (5.4)	59.8 (5.5)	60.3 (5.5)	59.7 (5.5)
Total excess MET-h/wk: mean (SD)	5.4 (2.8)	25.5 (10.9)	93.0 (45.9)	34.4 (33.1)
Vigorous excess MET-h/wk: mean (SD)	<1 (1.2)	6.4 (7.3)	22.8 (25.7)	8.9 (15.6)
Height, cm: mean (SD)	162.4 (6.2)	162.5 (6.1)	162.0 (6.1)	162.4 (6.1)
Weight, kg: mean (SD)	72.2 (13.6)	69.2 (11.9)	67.9 (11.4)	69.6 (12.2)
BMI, kg/m^2: mean (SD)	27.4 (5.0)	26.2 (4.3)	25.9 (4.1)	26.4 (4.5)
<18.50, %	<1	<1	<1	<1
18.5–24.99, %	34.5	42.7	46.1	41.6
25.00–29.99, %	39.6	39.7	38.4	39.3
≥30.00, %	25.4	16.9	14.8	18.2
Body fat percentage: mean (SD)	37.7 (6.4)	36.0 (6.3)	35.1 (6.4)	36.2 (6.4)
Body fat mass, kg: mean (SD)	28.0 (9.6)	25.5 (8.5)	24.4 (8.2)	25.8 (8.8)
Trunk fat percentage: mean (SD)	35.5 (7.2)	33.6 (7.2)	32.5 (7.4)	33.7 (7.3)
Trunk fat mass, kg: mean (SD)	14.3 (5.1)	13.0 (4.6)	12.4 (4.5)	13.1 (4.7)
Waist circumference, cm: mean (SD)	85.7 (12.0)	82.9 (10.9)	81.8 (10.7)	83.2 (11.2)
Hip circumference, cm: mean (SD)	104.2 (10.1)	101.9 (8.9)	100.8 (8.6)	102.1 (9.2)
Waist-to-hip ratio: mean (SD)	0.82 (0.07)	0.81 (0.07)	0.81 (0.07)	0.81 (0.07)
Qualifications[*] (%)				
College or university degree	33.1	34.4	27.0	32.4
National examination at ages 17–18[†] y	12.2	11.9	10.9	11.7
National examination at age 16[‡] y	29.1	27.5	28.7	28.1
Vocational qualifications	9.9	10.4	12.6	10.8
In paid employment or self-employed	59.2	48.0	43.9	49.5
Retired	35.7	46.1	49.4	44.6
Standing or walking job of all participants[§]	10.7	13.1	25.8	15.5
Manual job[‖] of all participants	1.3	2.5	9.8	3.9
Alcohol consumption (%)				
Never or rarely	20.2	18.3	21.6	19.5
Weekly	38.9	37.6	37.3	37.8
More than weekly	40.9	44.1	41.0	42.7
Smoking status (%)				
Never	58.8	59.7	59.7	59.5
Previous	32.8	33.6	33.3	33.4
Current	8.2	6.5	7.0	7.0
Parity: mean (SD)	1.8 (1.1)	1.9 (1.1)	1.9 (1.1)	1.9 (1.1)
Nulliparous, %	17.4	17.0	14.5	16.5
One or two, %	59.5	58.9	59.0	59.1
Three or more, %	23.1	24.0	26.4	24.4
Hormone replacement therapy use (%)				
Never	54.0	53.1	51.4	52.9
Previous	44.0	44.7	46.3	44.9
Current	1.9	2.0	2.2	2.0

PA = physical activity.
[*] Numbers do not add up to 100% due to missing data.
[†] Referred to as A levels, AS levels, or equivalent in UK Biobank.
[‡] Referred to as O levels, GCSEs, CSEs, or equivalent in UK Biobank.
[§] Participants who reported their work "usually" or "always" involved walking or standing for most of the time.
[‖] Participants who reported their work "usually" or "always" involved heavy manual or physical work for most of the time.

Age-adjusted Pearson's partial correlation coefficients for associations between different measures of body size and composition are shown in Tables 3 and 4. BMI was highly correlated with body fat mass in both premenopausal and postmenopausal women (0.94 and 0.93, respectively). BMI showed only a moderate correlation with WHR (0.46 and 0.44 in premenopausal and postmenopausal women, respectively). WHR was strongly correlated with waist circumference (0.75 in both) but weakly correlated with hip

Table 3
Age-adjusted Pearson's partial correlation coefficients for the 38,201 premenopausal women participants of UK Biobank

	Height, cm	Weight, kg	BMI, kg/m^2	Waist, cm	Hip, cm	WHR	Body fat %	Body fat mass, kg	Trunk fat %	Trunk fat mass, kg
Height, cm	1.00									
Weight, kg	0.30	1.00								
BMI, kg/m^2	−0.10	0.92	1.00							
Waist, cm	0.07	0.86	0.87	1.00						
Hip, cm	0.17	0.91	0.88	0.82	1.00					
WHR	−0.08	0.41	0.46	0.75	0.24	1.00				
Body fat %	0.05	0.84	0.86	0.82	0.82	0.45	1.00			
Body fat mass, kg	0.17	0.97	0.94	0.88	0.91	0.44	0.94	1.00		
Trunk fat %	0.16	0.81	0.78	0.77	0.79	0.40	0.98	0.90	1.00	
Trunk fat mass, kg	0.25	0.94	0.87	0.85	0.88	0.42	0.95	0.98	0.95	1.00

Table 4
Age-adjusted Pearson's partial correlation coefficients for the 94,592 postmenopausal women participants of UK Biobank

	Height, cm	Weight, kg	BMI, kg/m^2	Waist, cm	Hip, cm	WHR	Body fat %	Body fat mass, kg	Trunk fat %	Trunk fat mass, kg
Height, cm	1.00									
Weight, kg	0.31	1.00								
BMI, kg/m^2	−0.12	0.90	1.00							
Waist, cm	0.06	0.85	0.85	1.00						
Hip, cm	0.15	0.89	0.87	0.80	1.00					
WHR	−0.07	0.39	0.44	0.75	0.21	1.00				
Body fat %	0.02	0.82	0.84	0.79	0.80	0.42	1.00			
Body fat mass, kg	0.17	0.96	0.93	0.86	0.90	0.42	0.93	1.00		
Trunk fat %	0.14	0.78	0.75	0.73	0.75	0.38	0.98	0.90	1.00	
Trunk fat mass, kg	0.26	0.93	0.85	0.82	0.86	0.40	0.94	0.98	0.94	1.00

circumference (0.24 and 0.21 in premenopausal and postmenopausal women, respectively). Total and central adiposity were very highly correlated, as illustrated by the correlation coefficients of body with trunk fat mass (0.98 in both) and body with trunk fat percentage (0.98 in both).

Lower values for all body size and composition variables were associated with higher levels of physical activity in both premenopausal and postmenopausal women (Fig. 1). A test for the difference of nine means, for each anthropometric variable, detected a significant difference ($P < .0001$). The data for Figure 1 are shown in Supplementary Tables 3 and 4 for premenopausal and postmenopausal women, respectively. The mean BMI for premenopausal women who did less than 1 excess MET-h/wk was 27.9 kg/m^2 (95% CI: 27.6–28.2) compared with 25.0 kg/m^2 (95% CI: 24.9–25.2) for those who did 75 excess MET-h/wk or more of total physical activity. Mean BMI values for postmenopausal women were higher: 28.8 kg/m^2 (95% CI: 28.6–29.0) for those who did less than 1 MET-h/wk and 25.7 kg/m^2 (95% CI: 25.6–25.8) for those who did 75 excess MET-h/wk or more of total physical activity. For premenopausal women, the mean body fat percentage for those who did less than 1 excess MET-h/wk was 37.0% (95% CI: 36.5–37.4) compared with 31.9% (95% CI: 31.7–32.2) for those who did 75 excess MET-h/wk or more of total physical activity. Mean body fat percentage was higher for postmenopausal women: 39.4% (95% CI: 39.1–39.7) for those who did less than 1 MET-h/wk and 34.8% (95% CI: 34.6–34.9) for those who did 75 excess MET-h/wk or more of total physical activity. Figure 1 also shows similar patterns for the other anthropometric variables. The relationships were nonlinear with a greater absolute change in each measure of body size or composition for a given absolute change in physical activity at lower levels of physical activity. We ran separate analyses in manual workers and nonmanual workers and observed similar patterns of association between physical activity and adiposity (Supplementary Fig. 1).

Figure 2 shows the relationships between BMI, body fat percentage, WHR, and total physical activity according to the ratios of vigorous-to-total physical activity. The data for Figure 2 are shown in Supplementary Tables 5 and 6 for premenopausal and postmenopausal women, respectively. Regardless of menopausal status, within each category of total physical activity, a greater proportion of vigorous physical activity was associated with overall lower BMI, body fat percentage, and WHR ($P < .0001$ for all variables). The proportion of vigorous physical activity had a more distinct association with body composition at higher levels of total physical activity (>30 total excess MET-h/wk), where the lines diverge. For premenopausal women, at around 60 total excess MET-h/wk, those who did no vigorous physical activity had an average body fat percentage of 33.9% (95% CI: 33.3–34.6) compared with 31.6% (95% CI: 31.3–31.9) for those whose total physical activity consisted of more than 33% vigorous physical activity. The difference in body fat percentage from no vigorous physical activity to more than 33% vigorous physical activity was of slightly less magnitude in postmenopausal women: 36.0% (95% CI: 35.8–36.3) down to 34.5% (95% CI: 34.3–34.6) among those who did approximately 60 total excess MET-h/wk of physical activity.

Discussion

In this large cross-sectional study of 132,793 British women, we found that higher physical activity was associated with less adiposity across all measures for both premenopausal and postmenopausal women. Most women in our study, especially at lower levels of physical activity, were overweight, suggesting that it may be more difficult to maintain appropriate energy balance at low levels of physical activity [15]. The relationship between higher physical activity and lower adiposity was steeper for comparisons within the lower range; about more than 50 excess MET-h/wk, further increase in activity was associated with a relatively small change in measures of adiposity. This may point to a possible considerable benefit of physical activity for those who are relatively inactive, although it is possible that this pattern of risk may partly reflect reverse causality and/or measurement error, with overreporting of high physical activity.

Our results agree with previous cross-sectional studies that found a strong correlation between BMI and body fat percentage [16–18]. High age-adjusted correlation coefficients of BMI with weight, waist circumference, hip circumference, body fat percentage, body fat mass, trunk fat percentage, and trunk fat mass suggest that, within our study population, BMI is in fact a fairly good indicator of both total and central adiposity. WHR showed weaker correlation with BMI and the other previously mentioned measurements, with the exception of waist circumference. Accordingly, WHR has shown different associations with disease risk [19,20] that may capture relevant information beyond adiposity.

We found variation in physical activity by occupation. For instance, 5% of our participants held jobs that usually or always included manual work. Manual workers were characterized by much higher mean levels of both total and vigorous physical activity. They were also less likely to possess a college or university degree than the overall population. The overall means of the various measures of adiposity were similar between those who did and did not hold jobs that at least usually involved manual work. However, within the same level of total physical activity, manual workers tended to have a higher BMI than nonmanual workers. However, the pattern of association between higher physical activity and lower adiposity was similar when comparing manual with nonmanual workers.

Higher intensity exercise results in greater energy expenditure for a given duration. After taking into account total energy expenditure, some previous randomized controlled trials suggest that high-intensity exercise is more effective in reducing adiposity than low- or moderate-intensity exercise by improving control of energy balance [15,21]. However, other trials have found no additional

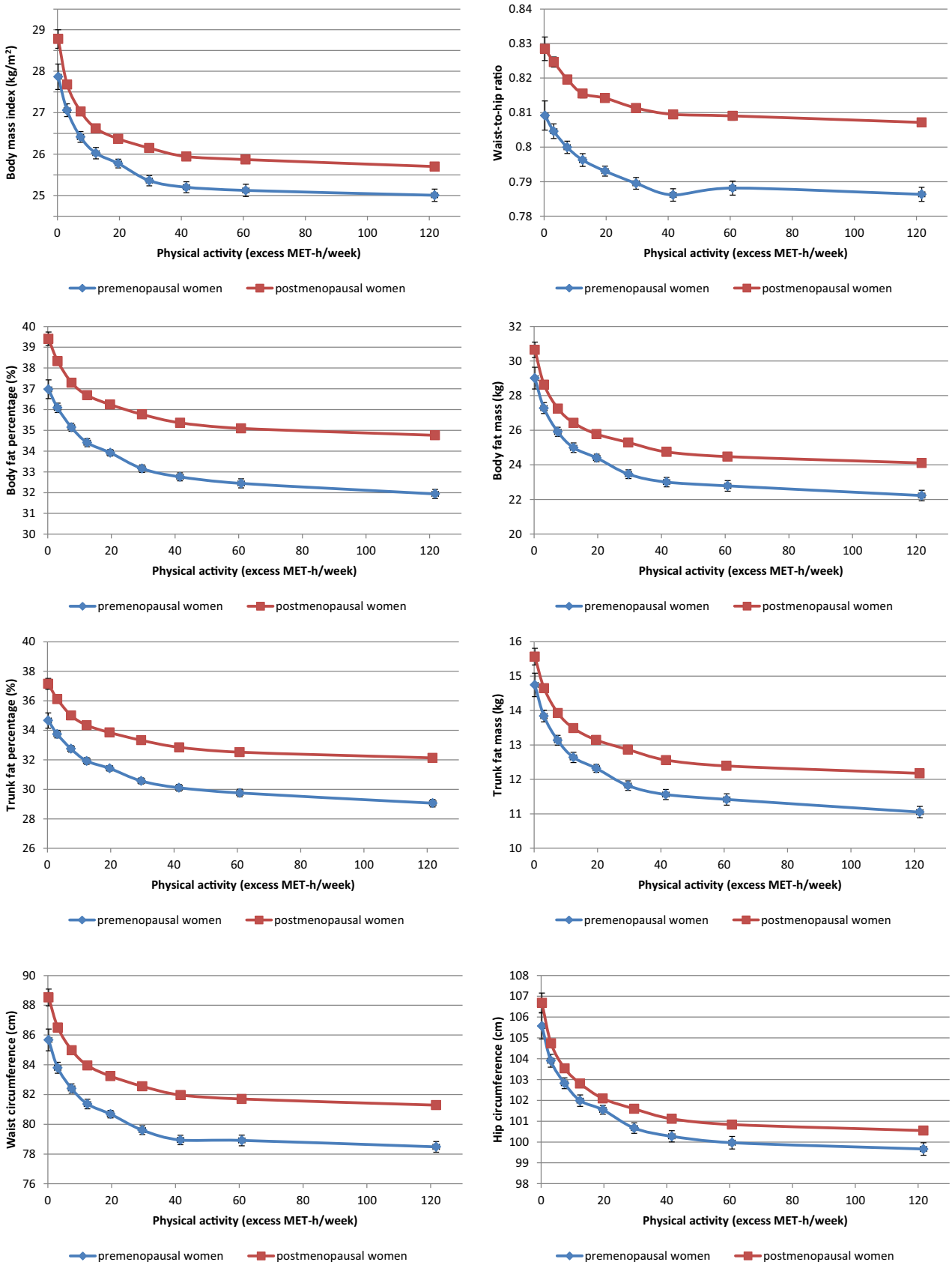


Fig. 1. Association between physical activity and body size and composition by menopausal status for the 38,201 premenopausal and 94,592 postmenopausal women of UK Biobank, 2006 to 2010. Physical activity was grouped in the following categories: 0, 1–4.9, 5–9.9, 10–14.9, 15–24.9, 25–34.9, 35–49.9, 50–74.9, and ≥75 excess MET-h/wk and are plotted at the value of the mean excess MET-h/wk in each category. Error bars are 95% confidence intervals. All results from the test for difference of nine means were highly significant ($P < .0001$).

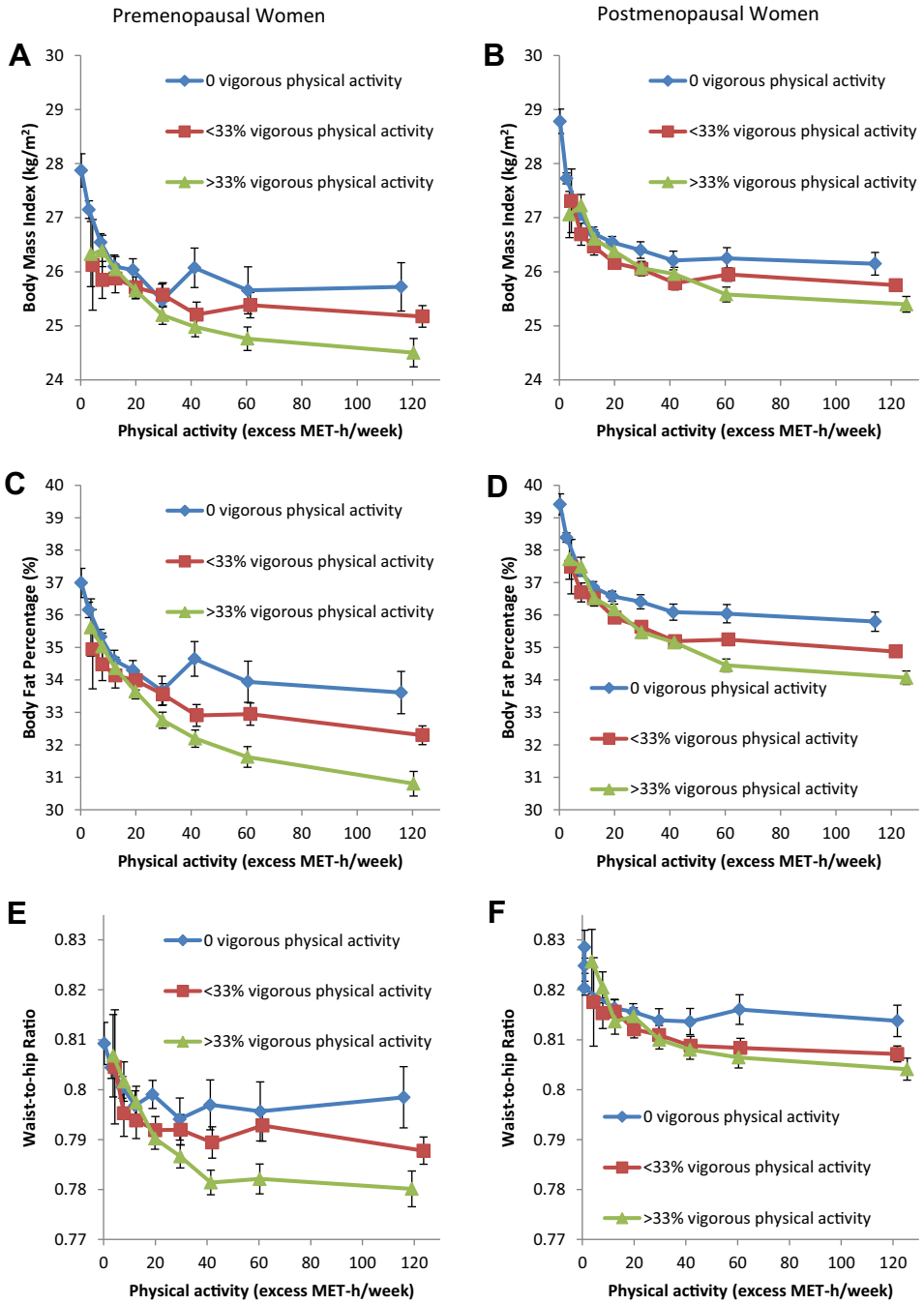


Fig. 2. Association between total physical activity and body size and composition by categories of vigorous physical activity for the 38,201 premenopausal and 94,592 postmenopausal women of UK Biobank, 2006 to 2010. Panels A, C, E correspond to premenopausal women, and panels B, D, and F correspond to postmenopausal women. Physical activity was grouped in the following categories: 0, 1–4.9, 5–9.9, 10–14.9, 15–24.9, 25–34.9, 35–49.9, 50–74.9, and ≥75 excess MET-h/wk and are plotted at the value of the mean excess MET-h/wk in each category. Error bars are 95% confidence intervals. Likelihood ratio tests were used to assess whether the association between total physical activity and the particular anthropometric outcome variable is significantly different depending on the different levels of vigorous-to-total physical activity. In all cases, a highly statistically significant difference was detected ($P < .0001$).

benefit of higher exercise intensity on abdominal fat loss [22–24] and weight loss [25,26]. Although differential measurement error is possible, our results suggest that a higher proportion of vigorous physical activity within each category of total physical activity was associated with less adiposity. Given the same energy expenditure, high-intensity exercise may stimulate greater fat oxidation and oxygen consumption during the postexercise period [27,28].

In agreement with previous studies, we found higher values for all body size variables in postmenopausal than in premenopausal women across all categories of physical activity. The menopausal transition results in cessation of ovarian function leading to subsequent decline in estrogen and progesterone levels, which has been associated with adverse changes in body composition and an increase in central adiposity [29,30]. Studies have reported similar total adiposity and fat distribution in premenopausal and hormone replacement therapy-treated postmenopausal women, suggesting an important role for ovarian hormones in determining fat accumulation and distribution [31,32]. Estrogen may protect against

increased adiposity through regulating appetite, resting metabolic rate, and levels of lipogenesis and lipoprotein lipase activity [33].

We are not aware of any other very large studies that have examined the association of total and intensity-specific physical activity with adiposity by menopausal status. A major strength of this study is the availability of a variety of objectively measured body size and composition measurements. Due to large study size, we were able to exclude women who reported a long-standing illness, disability, or infirmity, which might affect both energy balance and physical activity. We decided to use the IPAQ short form data processing rules because of lack of comprehensive coverage of domain-specific activities on the UK Biobank questionnaire. Although our study is limited by its lack of detailed information on associations between adiposity and levels of specific activities, we felt using questions that captured the frequency and duration of walking, moderate, and vigorous activities was the more straightforward and unbiased approach to estimate physical activity levels.

Limitations include the subjective reporting of physical activity at a single point in time, incomplete reporting of some variables, and lack of data on total energy intake. Energy intake and total energy expenditure are currently being assessed in participants using repeated web-based dietary questionnaires and the doubly labeled water method. In this cross-sectional analysis, we cannot assess whether differences in physical activity are causally related to varying levels of adiposity. For instance, those who are highly active may be individuals who are also more successful at maintaining appropriate target energy intake. In the future, UK Biobank accelerometer data will provide objective measures of physical activity frequency, duration, and intensity for many participants, and this information may help resolve some of the issues around self-reported physical activity. Health outcomes of all UK Biobank participants are being followed long-term, and future analyses will be able to include prospective studies of the relationships between physical activity, body size and composition, and disease outcome [34].

In conclusion, our analyses illustrate that UK women reporting more physical activity have lower levels of adiposity, as indicated by a range of body size and composition measures. Additionally, a higher proportion of vigorous physical activity was associated with even lower levels of adiposity. These results suggest that exercise intensity may be associated with body size and composition, beyond the influence of exercise frequency and duration.

Acknowledgments

This work was supported by Cancer Research UK, grant number C570/A11691. Wenji Guo is supported by the Clarendon Fund. The authors thank Paul Appleby for assistance with statistical analysis and Jane Green for comments on the article.

Supplementary data

Supplementary data related to this article can be found at http://dx.doi.org/10.1016/j.annepidem.2015.01.015.

References

[1] Must A, Spadano J, Coakley EH, Field AE, Colditz G, Dietz WH. The disease burden associated with overweight and obesity. JAMA 1999;282(16):1523–9.
[2] Britton KA, Lee IM, Wang L, Gaziano JM, Manson JE, Buring JE, et al. Physical activity and the risk of becoming overweight or obese in middle-aged and older women. Obesity (Silver Spring) 2012;20(5):1096–103.
[3] Ekelund U, Besson H, Luan J, May AM, Sharp SJ, Brage S, et al. Physical activity and gain in abdominal adiposity and body weight: prospective cohort study in 288,498 men and women. Am J Clin Nutr 2011;93(4):826–35.
[4] Catenacci VA, Wyatt HR. The role of physical activity in producing and maintaining weight loss. Nat Clin Pract Endocrinol Met 2007;3(7):518–29.
[5] Schwartz RS, Shuman WP, Larson V, Cain KC, Fellingham GW, Beard JC, et al. The effect of intensive endurance exercise training on body fat distribution in young and older men. Metabolism 1991;40(5):545–51.
[6] Kohrt WM, Obert KA, Holloszy JO. Exercise training improves fat distribution patterns in 60- to 70-year-old men and women. J Gerontol 1992;47(4):M99–105.
[7] Peters HW, Westendorp IC, Hak AE, Grobbee DE, Stehouwer CD, Hofman A, et al. Menopausal status and risk factors for cardiovascular disease. J Intern Med 1999;246(6):521–8.
[8] Key TJ, Verkasalo PK, Banks E. Epidemiology of breast cancer. Lancet Oncol 2001;2(3):133–40.
[9] Evenson KR, Wilcox S, Pettinger M, Brunner R, King AC, McTiernan A. Vigorous leisure activity through women's adult life: the Women's Health Initiative Observational Cohort Study. Am J Epidemiol 2002;156(10):945–53.
[10] UK Biobank. Protocol for a large-scale prospective epidemiological resource. 2010 [21 March 2014]; Available from: http://www.ukbiobank.ac.uk/wp-content/uploads/2011/11/UK-Biobank-Protocol.pdf.
[11] Craig CL, Marshall AL, Sjostrom M, Bauman AE, Booth ML, Ainsworth BE, et al. International physical activity questionnaire: 12-country reliability and validity. Med Sci Sports Exerc 2003;35(8):1381–95.
[12] Ainsworth BE, Haskell WL, Herrmann SD, Meckes N, Bassett Jr DR, Tudor-Locke C, et al. 2011 Compendium of Physical Activities: a second update of codes and MET values. Med Sci Sports Exerc 2011;43(8):1575–81.
[13] Armstrong ME, Cairns BJ, Green J, Reeves GK, Beral V. Reported frequency of physical activity in a large epidemiological study: relationship to specific activities and repeatability over time. BMC Med Res Methodol 2011;11:97.
[14] Morris R, Carstairs V. Which deprivation? A comparison of selected deprivation indexes. J Public Health Med 1991;13(4):318–26.
[15] Imbeault P, Saint-Pierre S, Almeras N, Tremblay A. Acute effects of exercise on energy intake and feeding behaviour. Br J Nutr 1997;77(4):511–21.
[16] Gray DS, Fujioka K. Use of relative weight and body mass index for the determination of adiposity. J Clin Epidemiol 1991;44(6):545–50.
[17] Neovius M, Linne Y, Rossner S. BMI, waist-circumference and waist-hip-ratio as diagnostic tests for fatness in adolescents. Int J Obes (Lond) 2005;29(2):163–9.
[18] Ranasinghe C, Gamage P, Katulanda P, Andraweera N, Thilakarathne S, Tharanga P. Relationship between Body Mass Index (BMI) and body fat percentage, estimated by bioelectrical impedance, in a group of Sri Lankan adults: a cross sectional study. BMC Public Health 2013;13:797.
[19] Amadou A, Hainaut P, Romieu I. Role of obesity in the risk of breast cancer: lessons from anthropometry. J Oncol 2013;2013:906495.
[20] Yusuf S, Hawken S, Ôunpuu S, Bautista L, Franzosi MG, Commerford P, et al. Obesity and the risk of myocardial infarction in 27,000 participants from 52 countries: a case-control study. Lancet 2005;366(9497):1640–9.
[21] Irving BA, Davis CK, Brock DW, Weltman JY, Swift D, Barrett EJ, et al. Effect of exercise training intensity on abdominal visceral fat and body composition. Med Sci Sports Exerc 2008;40(11):1863–72.
[22] Nicklas BJ, Wang X, You T, Lyles MF, Demons J, Easter L, et al. Effect of exercise intensity on abdominal fat loss during calorie restriction in overweight and obese postmenopausal women: a randomized, controlled trial. Am J Clin Nutr 2009;89(4):1043–52.
[23] Slentz CA, Aiken LB, Houmard JA, Bales CW, Johnson JL, Tanner CJ, et al. Inactivity, exercise, and visceral fat. STRRIDE: a randomized, controlled study of exercise intensity and amount. J Appl Physiol (1985) 2005;99(4):1613–8.
[24] Cho JK, Lee SH, Lee JY, Kang HS. Randomized controlled trial of training intensity in adiposity. Int J Sports Med 2011;32(6):468–75.
[25] Jakicic JM, Marcus BH, Gallagher KI, Napolitano M, Lang W. Effect of exercise duration and intensity on weight loss in overweight, sedentary women: a randomized trial. JAMA 2003;290(10):1323–30.
[26] Chambliss HO. Exercise duration and intensity in a weight-loss program. Clin J Sport Med 2005;15(2):113–5.
[27] Yoshioka M, Doucet E, St-Pierre S, Almeras N, Richard D, Labrie A, et al. Impact of high-intensity exercise on energy expenditure, lipid oxidation and body fatness. Int J Obes Relat Metab Disord 2001;25(3):332–9.
[28] Knab AM, Shanely RA, Corbin KD, Jin F, Sha W, Nieman DC. A 45-minute vigorous exercise bout increases metabolic rate for 14 hours. Med Sci Sports Exerc 2011;43(9):1643–8.
[29] Donato GB, Fuchs SC, Oppermann K, Bastos C, Spritzer PM. Association between menopause status and central adiposity measured at different cutoffs of waist circumference and waist-to-hip ratio. Menopause 2006;13(2):280–5.
[30] Toth MJ, Tchernof A, Sites CK, Poehlman ET. Effect of menopausal status on body composition and abdominal fat distribution. Int J Obes Relat Metab Disord 2000;24(2):226–31.
[31] Genazzani AR, Gambacciani M. Effect of climacteric transition and hormone replacement therapy on body weight and body fat distribution. Gynecol Endocrinol 2006;22(3):145–50.
[32] Haarbo J, Marslew U, Gotfredsen A, Christiansen C. Postmenopausal hormone replacement therapy prevents central distribution of body fat after menopause. Metabolism 1991;40(12):1323–6.
[33] Davis SR, Castelo-Branco C, Chedraui P, Lumsden MA, Nappi RE, Shah D, et al. Understanding weight gain at menopause. Climacteric 2012;15(5):419–29.
[34] Allen N, Sudlow C, Downey P, Peakman T, Danesh J, Elliott P, et al. UK Biobank: Current status and what it means for epidemiology. Health Policy Technology 2012;1(3):123–6.

Appendix

Supplementary Table 1
Physical activity questions

Questions	Potential responses
In a typical week, on how many days did you walk for at least 10 min at a time? (Include walking that you do at work, traveling to and from work, and for sport or leisure)	Number of days, "do not know," "unable to walk," or "prefer not to answer"
How many minutes did you usually spend walking on a typical day?	Quantity of minutes, "do not know" or "prefer not to answer"
In a typical week, on how many days did you do 10 min or more of moderate physical activities like carrying light loads, cycling at normal pace? (Do not include walking)	Number of days, "do not know" or "prefer not to answer"
How many minutes did you usually spend doing moderate activities on a typical day?	Quantity of minutes, "do not know" or "prefer not to answer"
In a typical week, how many days did you do 10 min or more of vigorous physical activity? (These are activities that make you sweat or breathe hard such as fast cycling, aerobics, heavy lifting)	Number of days, "do not know" or "prefer not to answer"
How many minutes did you usually spend doing vigorous activities on a typical day?	Quantity of minutes, "do not know" or "prefer not to answer"

Supplementary Table 2
Characteristics of manual and nonmanual workers for premenopausal and postmenopausal women

Characteristics	Premenopausal women		Postmenopausal women	
	Manual workers	Nonmanual workers	Manual workers	Nonmanual workers
Number of participants	2660	35,561	3732	90,921
Age, y: mean (SD)	45.3 (3.4)	45.6 (3.4)	56.3 (5.0)	60.0 (5.4)
Total excess MET-h/wk: mean (SD)	76.7 (68.3)	30.1 (31.9)	84.1 (74.2)	34.9 (37.4)
Vigorous excess MET-h/wk: mean (SD)	17.4 (29.4)	9.2 (14.0)	18.6 (35.0)	8.6 (15.5)
Height, cm: mean (SD)	163.5 (6.3)	164.5 (6.2)	161.8 (6.2)	161.8 (6.2)
Weight, kg: mean (SD)	69.8 (13.6)	69.7 (13.4)	69.6 (13.1)	69.6 (12.2)
BMI, kg/m^2: mean (SD)	26.1 (5.0)	25.8 (4.7)	26.6 (4.7)	26.4 (4.4)
Body fat percentage: mean (SD)	33.7 (7.2)	33.8 (6.9)	35.6 (6.7)	36.2 (6.4)
Body fat mass, kg: mean (SD)	24.4 (9.8)	24.4 (9.6)	25.5 (9.3)	25.8 (8.7)
Trunk fat percentage: mean (SD)	30.8 (8.2)	31.3 (7.9)	32.9 (7.7)	33.8 (7.3)
Trunk fat mass, kg: mean (SD)	12.1 (5.2)	12.3 (5.2)	12.8 (5.0)	13.1 (4.7)
Waist circumference, cm: mean (SD)	81.2 (11.8)	80.6 (11.3)	83.4 (11.8)	83.2 (11.1)
Hip circumference, cm: mean (SD)	101.6 (9.9)	101.4 (9.5)	102.0 (9.7)	102.1 (9.1)
WHR: mean (SD)	0.80 (0.07)	0.79 (0.06)	0.82 (0.07)	0.81 (0.07)
College or university degree, %	24.9	47.3	17.4	33.0

Supplementary Table 3
Data for Figure 1—premenopausal women

Body size/composition measure	Mean MET-h/wk	MET-h/wk categories	*N*	Mean of body size/composition measure	Lower 95% confidence interval	Upper 95% confidence interval
BMI, kg/m^2	0.3	0	859	27.9	27.6	28.2
	3.1	1–4.9	3432	27.1	26.9	27.2
	7.5	5–9.9	4826	26.4	26.3	26.5
	12.4	10–14.9	4425	26.0	25.9	26.2
	19.7	15–24.9	7350	25.8	25.7	25.9
	29.6	25–34.9	5278	25.4	25.2	25.5
	41.6	35–49.9	4619	25.2	25.1	25.3
	60.8	50–74.9	3644	25.1	25.0	25.3
	121.7	≥75	3673	25.0	24.9	25.2
WHR	0.3	0	859	0.809	0.805	0.813
	3.1	1–4.9	3433	0.805	0.803	0.807
	7.5	5–9.9	4829	0.800	0.798	0.802
	12.4	10–14.9	4424	0.796	0.794	0.798
	19.7	15–24.9	7354	0.793	0.792	0.795
	29.6	25–34.9	5279	0.790	0.788	0.791
	41.6	35–49.9	4623	0.786	0.784	0.788
	60.8	50–74.9	3646	0.788	0.786	0.790
	121.7	≥75	3673	0.786	0.784	0.788
Body fat percentage, %	0.3	0	851	37.0	36.5	37.4
	3.1	1–4.9	3396	36.1	35.9	36.3
	7.5	5–9.9	4791	35.1	35.0	35.3
	12.4	10–14.9	4389	34.4	34.2	34.6
	19.7	15–24.9	7281	33.9	33.8	34.1
	29.6	25–34.9	5232	33.2	33.0	33.3
	41.6	35–49.9	4569	32.8	32.6	33.0
	60.8	50–74.9	3611	32.4	32.2	32.7
	121.7	≥75	3642	31.9	31.7	32.2
Body fat mass, kg	0.3	0	851	29.0	28.4	29.6
	3.1	1–4.9	3395	27.3	27.0	27.6
	7.5	5–9.9	4790	25.9	25.6	26.2
	12.4	10–14.9	4389	25.0	24.7	25.3
	19.7	15–24.9	7280	24.4	24.2	24.6
	29.6	25–34.9	5232	23.5	23.2	23.7
	41.6	35–49.9	4569	23.0	22.7	23.3
	60.8	50–74.9	3611	22.8	22.5	23.1
	121.7	≥75	3639	22.2	21.9	22.5
Trunk fat percentage, %	0.3	0	850	34.7	34.1	35.2
	3.1	1–4.9	3395	33.7	33.5	34.0
	7.5	5–9.9	4786	32.7	32.5	33.0
	12.4	10–14.9	4386	31.9	31.7	32.2
	19.7	15–24.9	7275	31.4	31.2	31.6
	29.6	25–34.9	5228	30.6	30.4	30.8
	41.6	35–49.9	4565	30.1	29.9	30.3
	60.8	50–74.9	3608	29.8	29.5	30.0
	121.7	≥75	3641	29.1	28.8	29.3
Trunk fat mass, kg	0.3	0	850	14.7	14.4	15.1
	3.1	1–4.9	3395	13.8	13.7	14.0
	7.5	5–9.9	4785	13.1	13.0	13.3
	12.4	10–14.9	4386	12.6	12.5	12.8
	19.7	15–24.9	7274	12.3	12.2	12.4
	29.6	25–34.9	5227	11.8	11.7	12.0
	41.6	35–49.9	4565	11.6	11.4	11.7
	60.8	50–74.9	3608	11.4	11.3	11.6
	121.7	≥75	3641	11.1	10.9	11.2
Waist circumference, cm	0.3	0	859	85.7	84.9	86.4
	3.1	1–4.9	3433	83.8	83.4	84.2
	7.5	5–9.9	4829	82.4	82.1	82.7
	12.4	10–14.9	4424	81.4	81.0	81.7
	19.7	15–24.9	7354	80.7	80.4	80.9
	29.6	25–34.9	5279	79.6	79.3	79.9
	41.6	35–49.9	4623	78.9	78.6	79.3
	60.8	50–74.9	3646	78.9	78.6	79.3
	121.7	≥75	3673	78.5	78.1	78.8
Hip Circumference, cm	0.3	0	859	105.6	104.9	106.2
	3.1	1–4.9	3433	103.9	103.6	104.2
	7.5	5–9.9	4829	102.8	102.6	103.1
	12.4	10–14.9	4424	102.0	101.7	102.3
	19.7	15–24.9	7354	101.5	101.3	101.8
	29.6	25–34.9	5279	100.7	100.4	100.9
	41.6	35–49.9	4623	100.3	100.0	100.5
	60.8	50–74.9	3646	100.0	99.7	100.3
	121.7	≥75	3673	99.7	99.4	100.0

Supplementary Table 4
Data for Figure 1—postmenopausal women

Body size/composition measure	Mean MET-h/wk	MET-h/wk categories	*N*	Mean of body size/composition measure	Lower 95% confidence interval	Upper 95% confidence interval
BMI, kg/m^2	0.3	0	1422	28.8	28.6	29.0
	3.1	1–4.9	7591	27.7	27.6	27.8
	7.5	5–9.9	11,694	27.0	27.0	27.1
	12.4	10–14.9	10,621	26.6	26.5	26.7
	19.6	15–24.9	17,381	26.4	26.3	26.4
	29.6	25–34.9	12,267	26.1	26.1	26.2
	41.6	35–49.9	11,661	25.9	25.9	26.0
	60.7	50–74.9	10,211	25.9	25.8	26.0
	121.7	≥75	11,526	25.7	25.6	25.8
WHR	0.3	0	1424	0.828	0.825	0.832
	3.1	1–4.9	7590	0.825	0.823	0.826
	7.5	5–9.9	11,698	0.820	0.818	0.821
	12.4	10–14.9	10,628	0.816	0.814	0.817
	19.6	15–24.9	17,389	0.814	0.813	0.815
	29.6	25–34.9	12,272	0.811	0.810	0.813
	41.6	35–49.9	11,668	0.809	0.808	0.811
	60.7	50–74.9	10,217	0.809	0.808	0.810
	121.7	≥75	11,532	0.807	0.806	0.808
Body fat percentage, %	0.3	0	1407	39.4	39.1	39.7
	3.1	1–4.9	7514	38.3	38.2	38.5
	7.5	5–9.9	11,594	37.3	37.2	37.4
	12.4	10–14.9	10,534	36.7	36.6	36.8
	19.6	15–24.9	17,198	36.2	36.1	36.3
	29.6	25–34.9	12,173	35.8	35.7	35.9
	41.6	35–49.9	11,570	35.4	35.2	35.5
	60.7	50–74.9	10,116	35.1	35.0	35.2
	121.7	≥75	11,412	34.8	34.6	34.9
Body fat mass, kg	0.3	0	1407	30.6	30.2	31.1
	3.1	1–4.9	7513	28.6	28.4	28.8
	7.5	5–9.9	11,593	27.3	27.1	27.4
	12.4	10–14.9	10,533	26.4	26.3	26.6
	19.6	15–24.9	17,197	25.8	25.7	25.9
	29.6	25–34.9	12,169	25.3	25.1	25.4
	41.6	35–49.9	11,570	24.7	24.6	24.9
	60.7	50–74.9	10,114	24.5	24.3	24.6
	121.7	≥75	11,406	24.1	24.0	24.3
Trunk fat percentage, %	0.3	0	1406	37.2	36.8	37.5
	3.1	1–4.9	7508	36.1	35.9	36.3
	7.5	5–9.9	11,588	35.0	34.9	35.1
	12.4	10–14.9	10,528	34.3	34.2	34.5
	19.6	15–24.9	17,192	33.9	33.7	34.0
	29.6	25–34.9	12,164	33.3	33.2	33.5
	41.6	35–49.9	11,563	32.9	32.7	33.0
	60.7	50–74.9	10,114	32.5	32.4	32.7
	121.7	≥75	11,410	32.1	32.0	32.3
Trunk fat mass, kg	0.3	0	1406	15.6	15.3	15.8
	3.1	1–4.9	7508	14.6	14.5	14.8
	7.5	5–9.9	11,588	13.9	13.8	14.0
	12.4	10–14.9	10,527	13.5	13.4	13.6
	19.6	15–24.9	17,192	13.1	13.1	13.2
	29.6	25–34.9	12,164	12.9	12.8	12.9
	41.6	35–49.9	11,562	12.6	12.5	12.6
	60.7	50–74.9	10,114	12.4	12.3	12.5
	121.7	≥75	11,409	12.2	12.1	12.3
Waist circumference, cm	0.3	0	1424	88.5	88.0	89.1
	3.1	1–4.9	7591	86.5	86.3	86.8
	7.5	5–9.9	11,698	85.0	84.8	85.2
	12.4	10–14.9	10,628	84.0	83.7	84.2
	19.6	15–24.9	17,389	83.2	83.1	83.4
	29.6	25–34.9	12,273	82.6	82.4	82.7
	41.6	35–49.9	11,668	82.0	81.8	82.2
	60.7	50–74.9	10,217	81.7	81.5	81.9
	121.7	≥75	11,533	81.3	81.1	81.5
Hip circumference, cm	0.3	0	1424	106.7	106.2	107.2
	3.1	1–4.9	7590	104.7	104.5	104.9
	7.5	5–9.9	11,698	103.5	103.4	103.7
	12.4	10–14.9	10,628	102.8	102.6	103.0
	19.6	15–24.9	17,389	102.1	102.0	102.2
	29.6	25–34.9	12,272	101.6	101.4	101.8
	41.6	35–49.9	11,668	101.1	100.9	101.3
	60.7	50–74.9	10,217	100.8	100.7	101.0
	121.7	≥75	11,532	100.5	100.4	100.7

Supplementary Table 5
Data for Figure 2—premenopausal women

MET-h/wk category	Mean MET-h/wk	Vigorous PA, %	*N*	Mean body mass index, kg/m^2	Lower 95% confidence interval	Upper 95% confidence interval
0	0.3	0	859	27.9	27.6	28.2
1–4.9	3.0	0	3092	27.1	27.0	27.3
	4.3	<33	116	26.1	25.3	27.0
	3.7	>33	224	26.3	25.7	26.9
5–9.9	7.3	0	3165	26.5	26.4	26.7
	7.9	<33	694	25.8	25.5	26.2
	7.8	>33	967	26.4	26.1	26.7
10–14.9	12.2	0	1792	26.1	25.9	26.3
	12.6	<33	1169	25.9	25.6	26.1
	12.6	>33	1464	26.0	25.8	26.3
15–24.9	18.9	0	1939	26.0	25.8	26.2
	20.0	<33	2076	25.7	25.5	25.9
	19.9	>33	3335	25.7	25.5	25.8
25–34.9	29.4	0	860	25.5	25.2	25.8
	29.8	<33	1614	25.6	25.3	25.8
	29.6	>33	2804	25.2	25.0	25.4
35–49.9	41.2	0	608	26.1	25.7	26.4
	41.9	<33	1537	25.2	25.0	25.4
	41.4	>33	2474	25.0	24.8	25.2
50–74.9	60.5	0	432	25.7	25.2	26.1
	61.4	<33	1504	25.4	25.2	25.6
	60.3	>33	1708	24.8	24.5	25.0
≥75	115.9	0	408	25.7	25.3	26.2
	123.6	<33	2062	25.2	25.0	25.4
	120.4	>33	1203	24.5	24.2	24.8

MET-h/wk category	Mean MET-h/wk	Vigorous PA, %	*N*	Mean body fat percentage, %	Lower 95% confidence interval	Upper 95% confidence interval
0	0.3	0	851	37.0	36.5	37.4
1–4.9	3.0	0	3061	36.2	35.9	36.4
	4.3	<33	115	34.9	33.7	36.2
	3.7	>33	220	35.6	34.7	36.5
5–9.9	7.3	0	3139	35.3	35.1	35.6
	7.9	<33	690	34.5	34.0	35.0
	7.8	>33	962	35.0	34.6	35.4
10–14.9	12.2	0	1779	34.6	34.3	34.9
	12.6	<33	1156	34.1	33.8	34.5
	12.6	>33	1454	34.4	34.0	34.7
15–24.9	18.9	0	1927	34.3	34.0	34.6
	20.0	<33	2054	34.0	33.7	34.3
	19.9	>33	3300	33.6	33.4	33.9
25–34.9	29.4	0	851	33.7	33.2	34.1
	29.8	<33	1602	33.6	33.2	33.9
	29.6	>33	2779	32.8	32.5	33.0
35–49.9	41.2	0	602	34.6	34.1	35.2
	41.9	<33	1522	32.9	32.6	33.2
	41.4	>33	2445	32.2	31.9	32.5
50–74.9	60.5	0	429	33.9	33.3	34.6
	61.4	<33	1483	32.9	32.6	33.3
	60.3	>33	1699	31.6	31.3	31.9
≥75	115.9	0	408	33.6	33.0	34.3
	123.6	<33	2042	32.3	32.0	32.6
	120.4	>33	1192	30.8	30.4	31.2

MET-h/wk category	Mean MET-h/wk	Vigorous PA, %	*N*	Mean WHR	Lower 95% confidence interval	Upper 95% confidence interval
0	0.3	0	859	0.809	0.805	0.813
1–4.9	3.0	0	3093	0.804	0.802	0.807
	4.3	<33	116	0.805	0.793	0.816
	3.7	>33	224	0.807	0.799	0.815
5–9.9	7.3	0	3167	0.800	0.798	0.803
	7.9	<33	695	0.795	0.791	0.800
	7.8	>33	967	0.802	0.798	0.806
10–14.9	12.2	0	1792	0.797	0.794	0.800
	12.6	<33	1169	0.794	0.790	0.797
	12.6	>33	1463	0.797	0.794	0.801
15–24.9	18.9	0	1940	0.799	0.796	0.802
	20.0	<33	2077	0.792	0.789	0.795
	19.9	>33	3337	0.790	0.788	0.792
25–34.9	29.4	0	860	0.794	0.790	0.798
	29.8	<33	1614	0.792	0.789	0.795
	29.6	>33	2805	0.787	0.784	0.789

(*continued on next page*)

Supplementary Table 5 (*continued*)

MET-h/wk category	Mean MET-h/wk	Vigorous PA, %	*N*	Mean WHR	Lower 95% confidence interval	Upper 95% confidence interval
35–49.9	41.2	0	608	0.797	0.792	0.802
	41.9	<33	1539	0.789	0.786	0.793
	41.4	>33	2476	0.781	0.779	0.784
50–74.9	60.5	0	433	0.796	0.790	0.802
	61.4	<33	1505	0.793	0.790	0.796
	60.3	>33	1708	0.782	0.779	0.785
≥75	115.9	0	408	0.798	0.792	0.805
	123.6	<33	2062	0.788	0.785	0.791
	119.1	>33	1203	0.780	0.777	0.784

PA = physical activity.

Supplementary Table 6
Data for Figure 2—postmenopausal women

MET-h/wk category	Mean MET-h/wk	Vigorous PA, %	*N*	Mean BMI, kg/m^2	Lower 95% confidence interval	Upper 95% confidence interval
0	0.3	0	1422	28.8	28.6	29.0
1–4.9	2.6	0	6991	27.7	27.6	27.8
	4.3	<33	207	27.3	26.7	27.9
	3.6	>33	393	27.1	26.6	27.5
5–9.9	7.4	0	8248	27.1	27.0	27.2
	7.9	<33	1736	26.7	26.5	26.9
	7.8	>33	1710	27.2	27.0	27.4
10–14.9	12.3	0	5163	26.7	26.6	26.8
	12.5	<33	2912	26.5	26.3	26.6
	12.6	>33	2546	26.6	26.4	26.8
15–24.9	19.1	0	6211	26.5	26.4	26.7
	20.0	<33	5479	26.2	26.0	26.3
	19.8	>33	5691	26.4	26.3	26.5
25–34.9	29.4	0	3086	26.4	26.2	26.6
	29.7	<33	4442	26.1	25.9	26.2
	29.7	>33	4739	26.1	25.9	26.2
35–49.9	41.2	0	2363	26.2	26.0	26.4
	41.8	<33	4731	25.8	25.7	25.9
	41.6	>33	4567	26.0	25.8	26.1
50–74.9	60.6	0	1850	26.2	26.1	26.4
	61.1	<33	4592	26.0	25.8	26.1
	60.4	>33	3769	25.6	25.4	25.7
≥75	114.2	0	1664	26.1	25.9	26.4
	121.7	<33	6448	25.8	25.6	25.9
	125.5	>33	3414	25.4	25.3	25.5

MET-h/wk category	Mean MET-h/wk	Vigorous PA, %	*N*	Mean body fat percentage, %	Lower 95% confidence interval	Upper 95% confidence interval
0	0.3	0	1407	39.4	39.1	39.7
1–4.9	2.6	0	6915	38.4	38.2	38.5
	4.3	<33	207	37.5	36.7	38.3
	3.6	>33	392	37.7	37.1	38.3
5–9.9	7.4	0	8173	37.4	37.2	37.5
	7.9	<33	1722	36.7	36.4	37.0
	7.8	>33	1699	37.5	37.2	37.8
10–14.9	12.3	0	5119	36.9	36.7	37.0
	12.5	<33	2889	36.5	36.3	36.7
	12.6	>33	2526	36.5	36.3	36.8
15–24.9	19.1	0	6135	36.6	36.4	36.7
	20.0	<33	5431	35.9	35.8	36.1
	19.8	>33	5632	36.2	36.0	36.3
25–34.9	29.4	0	3051	36.4	36.2	36.6
	29.7	<33	4422	35.6	35.5	35.8
	29.7	>33	4700	35.5	35.3	35.6
35–49.9	41.2	0	2337	36.1	35.8	36.3
	41.8	<33	4695	35.2	35.0	35.4
	41.6	>33	4538	35.2	35.0	35.3
50–74.9	60.6	0	1823	36.0	35.8	36.3
	61.1	<33	4550	35.2	35.1	35.4
	60.4	>33	3743	34.5	34.3	34.6

(*continued on next page*)

Supplementary Table 6 (*continued*)

MET-h/wk category	Mean MET-h/wk	Vigorous PA, %	*N*	Mean body fat percentage, %	Lower 95% confidence interval	Upper 95% confidence interval
≥75	114.2	0	1645	35.8	35.5	36.1
	121.7	<33	6388	34.9	34.7	35.0
	125.5	>33	3379	34.1	33.9	34.3
MET-h/wk category	Mean MET-h/wk	Vigorous PA, %	*N*	Mean WHR	Lower 95% confidence interval	Upper 95% confidence interval
0	0.3	0	1424	0.829	0.825	0.832
1–4.9	2.6	0	6990	0.825	0.823	0.826
	4.3	<33	207	0.818	0.809	0.826
	3.6	>33	393	0.826	0.819	0.832
5–9.9	7.4	0	8251	0.820	0.819	0.822
	7.9	<33	1737	0.815	0.812	0.818
	7.8	>33	1710	0.820	0.817	0.824
10–14.9	12.3	0	5169	0.816	0.815	0.818
	12.5	<33	2912	0.816	0.813	0.818
	12.6	>33	2547	0.814	0.811	0.816
15–24.9	19.1	0	6216	0.816	0.814	0.817
	20.0	<33	5480	0.812	0.810	0.814
	19.8	>33	5693	0.815	0.813	0.816
25–34.9	29.4	0	3085	0.814	0.812	0.816
	29.7	<33	4447	0.811	0.809	0.813
	29.7	>33	4740	0.810	0.808	0.812
35–49.9	41.2	0	2365	0.814	0.811	0.816
	41.8	<33	4734	0.809	0.807	0.811
	41.6	>33	4569	0.808	0.806	0.810
50–74.9	60.6	0	1851	0.816	0.813	0.819
	61.1	<33	4594	0.808	0.806	0.810
	60.4	>33	3772	0.806	0.804	0.809
≥75	114.2	0	1664	0.814	0.811	0.817
	121.7	<33	6451	0.807	0.806	0.809
	125.5	>33	3417	0.804	0.802	0.806

PA = physical activity.

C.2 Adiposity and breast cancer risk in postmenopausal women: Results from the UK Biobank prospective cohort (International Journal of Cancer, 2018)

I am the first author of this paper, which is based on work in Chapter 8. I conducted the statistical analyses, drafted the paper, and revised it to address reviewers' comments, with input from the co-authors.

IJC
International Journal of Cancer

Adiposity and breast cancer risk in postmenopausal women: Results from the UK Biobank prospective cohort

Wenji Guo, Timothy J. Key and Gillian K. Reeves

Cancer Epidemiology Unit, Nuffield Department of Population Health, University of Oxford, Oxford, United Kingdom

Body size is an important modifiable risk factor for postmenopausal breast cancer. However, it remains unclear whether direct measures of fat mass are better indicators of risk than anthropometric measures, or whether central adiposity may contribute to risk beyond overall adiposity. We analyzed data from 162,691 postmenopausal women in UK Biobank followed from 2006 to 2014. Body size was measured by trained technicians. Multivariable-adjusted Cox regression was used to estimate relative risks. Analyses were stratified by age at recruitment, region and socioeconomic status, and adjusted for family history of breast cancer, age at menarche, age at first birth, parity, age at menopause, previous hormone replacement therapy use, smoking, alcohol intake, height, physical activity and ethnicity. We observed 2,913 incident invasive breast cancers during a mean 5.7 years of follow-up. There was a continuous increase in risk of postmenopausal breast cancer with increasing adiposity, across all measures. The point estimate, comparing women in the top (median 37.6 kg) to bottom (median 17.6 kg) quartile of body fat mass was 1.70 (95% confidence interval 1.52–1.90). The magnitudes of the associations between per SD increase in BMI and body fat mass with breast cancer risk were similar, suggesting impedance measures of fat were not substantially better indicators of risk than anthropometric measures. After adjusting for body fat mass, the associations between anthropometric measures of central adiposity and breast cancer risk were attenuated. The magnitude of risk, across all measures of adiposity, was greater in women who had been postmenopausal for 12 or more years.

Introduction

Although there are several important known risk factors for breast cancer, excess adiposity is one of the few modifiable ones. Previous studies have demonstrated an association between obesity and increased risk for postmenopausal breast cancer,[1] but questions remain about this relationship and its underlying mechanisms. Since adipose tissue is the major site of estrogen synthesis in postmenopausal women, the most prominent hypothesis is that more body fat is linked to increased postmenopausal breast cancer risk through estrogen-stimulated carcinogenesis.[2,3] For this reason, fat mass may be a better predictor of postmenopausal breast cancer risk than other measures of adiposity.

Due to high cost and difficulty of implementation, most large prospective cohort studies have been unable to directly address this question, instead examining the association between obesity and breast cancer risk using only traditional anthropometric measures such as body mass index (BMI), waist circumference and waist-to-hip ratio, which do not directly distinguish between lean and fat mass.[4–7] While BMI is unable to capture changes in adiposity with age, bioelectrical impedance measures reflect increases in fat mass with age.[8] Furthermore, it is unclear whether the distribution of adipose tissue is relevant for breast cancer risk.[9] Most studies rely on self-reported measures of body size, which tend to underestimate adiposity.[10] We aim here to clarify the association between obesity and postmenopausal breast cancer risk using data on bioelectrical impedance measures of body fat in UK Biobank, a nationwide study of 500,000 individuals.

Methods

Data source

Data were obtained from UK Biobank (reference number 3248, approved August 2013). Details of the rationale, design and survey methods for UK Biobank have been described elsewhere[11] and information on data available and access procedures are given on the study website (http://www.

Key words: breast cancer, adiposity, obesity, bioelectrical impedance, UK Biobank

Abbreviations: BMI: body mass index; CI: confidence interval; DXA: dual-energy X-ray absorptiometry; HRT: hormone replacement therapy; RR: relative risk

Additional Supporting Information may be found in the online version of this article.

Conflict of interest: The authors have declared no conflicts of interest.

Grant sponsor: Cancer Research UK; **Grant number:** C570/A16491; **Grant sponsor:** Clarendon Fund

DOI: 10.1002/ijc.31394

This is an open access article under the terms of the Creative Commons Attribution License, which permits use, distribution and reproduction in any medium, provided the original work is properly cited.

History: Received 3 Jan 2018; Accepted 13 Mar 2018; Online 23 Mar 2018

Correspondence to: Wenji Guo, Cancer Epidemiology Unit, Nuffield Department of Population Health, University of Oxford, Richard Doll Building, Roosevelt Drive, Oxford OX3 7LF, United Kingdom, Tel.: 44-1865-289635, E-mail: wenji.guo@ndph.ox.ac.uk

What's new?

Obesity is associated with an increased risk of postmenopausal breast cancer. However, whether direct anaylses of fat mass using bioelectrical impedance measures are better indicators of breast cancer risk than anthropometric measures, such as body mass index (BMI) and waist circumference, remains unclear. In this study, based on data for postmenopausal women obtained from UK Biobank, the authors found that BMI, waist circumference, and direct bioimpedance measures of fat showed similarly strong associations with postmenopausal breast cancer risk. Central adiposity was no longer associated with breast cancer risk after adjusting for overall body fat mass.

ukbiobank.ac.uk/). UK Biobank has approval from the North West Multi-centre Research Ethics Committee, the Confidentiality Advisory Group in England and Wales and the Community Health Index Advisory Group in Scotland. All participants provided written informed consent.

Study participants

The complete UK Biobank dataset includes 502,620 UK adults (229,165 men and 273,455 women) aged 40–70 at recruitment during 2006–2010. Participants completed a touchscreen questionnaire during the baseline assessment center visit that included questions on socio-demographics, lifestyle, health and medical history and sex-specific factors. A number of physical measurements, including body size and composition, were also assessed on the whole cohort during the baseline assessment center visit. A repeat assessment of all baseline measures was conducted in 20,345 participants at the UK Biobank Coordinating Centre between August 2012 and June 2013.

Women were eligible for these analyses if they were postmenopausal at recruitment. Women were defined as being postmenopausal at recruitment if they reported that their periods had stopped. Women with unknown self-reported menopausal status were defined as postmenopausal if they were aged 53 or over at recruitment based on previously established criteria[12] and because >97% of the study population with known menopausal status reported having become postmenopausal by that age. Women were excluded if they had a prior cancer diagnosis (except for non-melanoma skin cancer ICD-10 C44) ($n = 18{,}372$), were premenopausal ($n = 62{,}899$) or had unknown menopausal status after applying the above categorization criteria ($n = 12{,}563$), were current users of hormone replacement therapy (HRT) ($n = 13{,}680$), or had missing data on any measure of body size and composition ($n = 3{,}250$). Current HRT users were excluded because HRT use is known to attenuate associations of adiposity with breast cancer risk in postmenopausal women.[4,13,14] 162,691 postmenopausal women were included in the analyses (Supporting Information Fig. S1).

Anthropometry and body composition

At the UK Biobank baseline interview, trained staff measured standing height using the Seca 202 device (Seca, Hamburg, Germany). The Wessex non-stretchable sprung tape measure (Wessex, United Kingdom) was used to measure waist circumference and hip circumference, from which we derived waist-to-hip-ratio by dividing waist circumference by hip circumference. The Tanita BC-418MA body composition analyzer (Tanita, Tokyo, Japan) was used to measure body and trunk fat mass and percentage using bio-impedance. BMI was calculated by dividing weight (kg) by the square of standing height (m^2). Dual-energy X-ray absorptiometry (DXA) was used to measure fat mass/percentage on a subset of participants beginning in 2014 using the GE-Lunar iDXA (GE Healthcare, Chicago, IL).

Ascertainment of cancer cases

UK Biobank obtains data on cancer diagnoses through the Health & Social Care Information Centre for participants in England and Wales, and the NHS Central Register for participants in Scotland.

Statistical analyses

Baseline characteristics of study participants were compared between top and bottom quartiles of BMI, body fat mass and waist circumference. Associations between the various measures of body size and composition were examined by calculating age-adjusted Pearson's partial correlation coefficients.

Women were followed from date of baseline assessment center visit until the earliest of: date of breast cancer registration (ICD-10 C50), date of death, date of loss to follow-up or end of follow-up for cancer incidence (November 30, 2014). Women diagnosed with any cancer other than breast cancer (with the exception of non-melanoma skin cancer) during follow-up were censored at date of diagnosis.

Multivariable-adjusted Cox regression with attained age as the underlying time variable was used to estimate hazards ratios (referred to as relative risks [RRs]) and 95% confidence intervals (CIs) for the association between breast cancer risk and body size and composition measures: height, weight, BMI, body fat mass, body fat percentage, waist circumference, hip circumference, waist-to-hip ratio, trunk fat mass and trunk fat percentage. For the risk analyses, women were categorized into quartiles according to their baseline body size/composition. To facilitate comparison across measures of adiposity, we also estimated RRs per standard deviation (SD) increase and show the results from tests for linear trend corrected using the repeat assessment median within each category (grouped into quartiles) and the corresponding χ^2 statistic. Tests for linear trend were performed with categories

coded in an ordinal fashion using the repeat assessment median of each category to attain measures that are more representative of the true long-term values.[15] In a separate analysis, tests for linear trend were performed using the DXA median of each quartile for fat mass/percentage variables.

All analyses were stratified by 5-year age at recruitment categories, region of recruitment and socioeconomic status (based on quintiles of Townsend deprivation index),[16] which allows the hazard function to vary across levels of the stratification variables. All analyses were adjusted for family history of breast cancer (no, yes), age at menarche (<12, 12–13, ≥14), age at first birth (<25, 25–29, ≥30), parity (nulliparous, 1–2, ≥3), age at menopause (<45, 45–54, ≥55), previous HRT use (never, past), smoking (never, past, current), alcohol intake frequency (<3 times a month, 1–4 times a week, daily or almost daily), physical activity (<14, 14–30, 31–59.8, ≥59.9 metabolic equivalent hours per week) and ethnicity (White, Mixed, Asian or Asian British, Black or Black British, Chinese, other ethnic group). Regression models with exposures other than height as the exposure of interest were adjusted for height (<160, 160–165.9, ≥166 cm) as an independent risk factor for breast cancer. Analyses were further adjusted for body size at age 10 (thinner, about average, plumper) to account for the independent association between greater childhood body fatness and decreased breast cancer risk.[17] Women with missing values for any of the adjustment variables were assigned to a separate "unknown" category for the respective variable. Information was either missing or reported as unknown for <3% of covariates, with the exception of total physical activity (28.5%) and age at menopause (18.7%) which may be difficult to accurately report if menopausal status is masked by hysterectomy or HRT before menstrual periods stop naturally.[12] To assess the impact of missing values, we conducted a sensitivity analysis restricted to participants with known values for all adjustment variables.

The χ^2 statistic for trend across quartile medians was used to quantify the extent to which each particular index of adiposity was related to risk.[18] We further adjusted anthropometric measures of central adiposity by the impedance measure of body fat mass, but we chose not to mutually adjust impedance measures for each other, because these variables were highly correlated (r >0.88). Assessment of interaction terms between each exposure of interest and the underlying time variable did not suggest any significant deviation from proportional hazards.

RRs were also estimated for each body size/composition variable as a function of years since menopause (<12, ≥12) by using a time-varying covariate. Comparisons were made between women who were <12 years and ≥12 years since menopause because this cut point divided the number of cases approximately equally. Likelihood ratio tests were used to assess whether time since menopause was an effect modifier in the associations between body size/composition variables and breast cancer risk. All analyses were conducted using STATA version 15.0 (Stata Corp LP, College Station, TX).

Results

Two thousand nine hundred and thirteen invasive breast cancer cases were diagnosed among 162,691 postmenopausal women at risk during a mean follow-up of 5.7 (SD 1.1) years. Table 1 shows baseline characteristics of participants according to BMI, body fat mass and waist circumference. Women in the top quartiles for these body size/composition variables were more likely to be in the lowest fifth of socioeconomic class, have a younger age at first birth, have used HRT, consume alcohol only occasionally or never and engage in less physical activity while the leanest women were more likely to have used oral contraceptives in the past. Those in the bottom quartiles of BMI and body fat mass were more likely to be either never or current smokers.

Table 2 shows age-adjusted Pearson's partial correlation coefficients for the various anthropometric and body composition measures. BMI was highly correlated with both body fat percentage (0.85) and body fat mass (0.94). Among the indicators of central obesity, trunk fat percentage and mass were highly correlated with waist circumference (0.74 and 0.83 respectively) and hip circumference (0.75 and 0.86 respectively) but not with waist-to-hip ratio (0.39 and 0.41 respectively). Waist-to-hip ratio was much more strongly correlated with waist circumference (0.74) than hip circumference (0.23). Total and central adiposity were very highly correlated as shown by the correlation coefficients between body and trunk fat percentage (0.98) and body and trunk fat mass (0.97). In view of these high correlations between certain indices, associations of risk with overall fatness were not mutually adjusted for other indices of body fatness. We present results for both body and trunk fat mass, but the extremely high correlation coefficients observed between these two indices indicate that their effects may be indistinguishable.

All measures of overall adiposity, as well as height, were associated with increased postmenopausal breast cancer risk in an approximately linear fashion (Fig. 1, Table 3). After adjusting for body size at age 10, the magnitudes of association for all indices of adiposity were increased. Lowest breast cancer risk was associated with being in the lowest quartile of adiposity (median BMI 22.2 kg/m^2). Women in the top quartile of body fat mass (32.6–108.4 kg; median, 37.6 kg), had the greatest increase in breast cancer risk (RR, 1.70; 95% CI, 1.52–1.90; p_{trend} <0.001), compared to women in the lowest quartile (8.0–20.5 kg; median, 17.6 kg). The magnitudes of associations with breast cancer risk per SD increases in BMI and body fat mass were similar with overlapping confidence intervals: RR 1.21; 95% CI 1.15–1.27 for BMI and RR 1.25; 95% CI 1.19–1.31 for body fat mass in multivariable-adjusted models (Tables 3 and 4). Recoding quartiles using the corresponding DXA instead of baseline medians for tests of linear trend produced nearly identical results (Supporting Information Table S1).

Table 1. Baseline characteristics for the 162,691 postmenopausal participants of UK Biobank by various measures of adiposity

	Body mass index (kg/m^2)		Body fat mass (kg)		Waist circumference (cm)	
	Quartile 1	Quartile 4	Quartile 1	Quartile 4	Quartile 1	Quartile 4
Age at recruitment, mean (SD)	59.8 (5.4)	60.3 (5.4)	59.9 (5.4)	60.3 (5.4)	59.6 (5.4)	60.6 (5.3)
Lowest socioeconomic quintile, %	16	25.1	16.7	23.8	15.2	25.5
Family history of breast cancer, %	7.2	6.1	6.8	6.3	7	6.2
Age at menarche, mean (SD)	13.2 (1.6)	12.7 (1.7)	13.1 (1.6)	12.7 (1.7)	13.1 (1.6)	12.7 (1.7)
Age at first birth, mean (SD)	26.4 (4.9)	24.6 (4.8)	26.2 (4.9)	24.7 (4.9)	26.2 (4.9)	24.7 (4.9)
Parity, mean (SD)	1.8 (1.2)	2.0 (1.3)	1.8 (1.2)	2.0 (1.2)	1.8 (1.1)	2.0 (1.3)
Ever oral contraceptive use, %	78.8	75.3	78.3	76.2	79	75.1
Age at menopause, mean (SD)	49.9 (4.6)	49.6 (5.5)	49.9 (4.7)	49.6 (5.4)	49.9 (4.6)	49.6 (5.4)
Past hormone replacement therapy use, %	43.1	47.6	42.6	47.8	42.7	47.5
Smoking, %						
Never	61	56.5	61.1	55.7	62.6	54.5
Past	29.7	35.6	29.3	36.5	29.1	36.6
Current	9.1	7.4	9.4	7.2	8	8.4
Alcohol intake frequency, %						
Special occasions only or never	21.1	36.2	22.2	34.4	21.3	35.8
1–3 times a month	10.3	15.1	10.6	14.8	11	14.4
1–4 times a week	46.3	38.4	45.8	39.6	47.5	38
Daily or almost daily	22.2	10.3	21.3	11.2	20.1	11.7
Height (cm), mean (SD)	162.9 (6.2)	160.8 (6.1)	160.5 (6.2)	162.9 (6.1)	161.5 (6.1)	162.0 (6.2)
Total physical activity (MET-hr), mean (SD)	53.3 (56.8)	40.3 (50.2)	55.1 (58.7)	39.0 (48.4)	53.9 (57.0)	39.6 (49.3)

Table 2. Age-adjusted Pearson's partial correlation coefficients for the 162,691 postmenopausal participants

	Height, cm	Weight, kg	BMI, kg/m^2	Waist circumference, cm	Hip circumference, cm	Waist-to-hip ratio	Body fat mass, kg	Body fat %	Trunk fat mass, kg	Trunk fat %
Height, cm	1.00									
Weight, kg	0.27	1.00								
BMI, kg/m^2	−0.13	0.92	1.00							
Waist circumference, cm	0.04	0.86	0.87	1.00						
Hip circumference, cm	0.12	0.91	0.89	0.82	1.00					
Waist-to-hip ratio	−0.07	0.41	0.45	0.74	0.23	1.00				
Body fat mass, kg	0.15	0.97	0.94	0.88	0.91	0.43	1.00			
Body fat %	0.01	0.83	0.85	0.81	0.81	0.44	0.93	1.00		
Trunk fat mass, kg	0.24	0.92	0.85	0.83	0.86	0.41	0.97	0.94	1.00	
Trunk fat %	0.13	0.77	0.75	0.74	0.75	0.39	0.88	0.98	0.94	1.00

All indicators of central obesity were significantly associated with postmenopausal breast cancer risk (Fig. 1, Table 4); fat mass was slightly more informative about postmenopausal breast cancer risk ($\chi^2 = 259.4$), and waist-to-hip ratio was least informative ($\chi^2 = 202.0$). After adjusting for body fat mass, associations of breast cancer risk with per SD increases in waist circumference, hip circumference and waist-to-hip ratio were attenuated to the null. However, the confidence intervals around the adjusted associations were moderately wide so small residual associations cannot be ruled out. Results using standard categories for BMI and waist circumference[19] are provided in Supporting Information Table S2. The results of sensitivity analyses excluding participants with any missing values did not differ materially from the main findings.

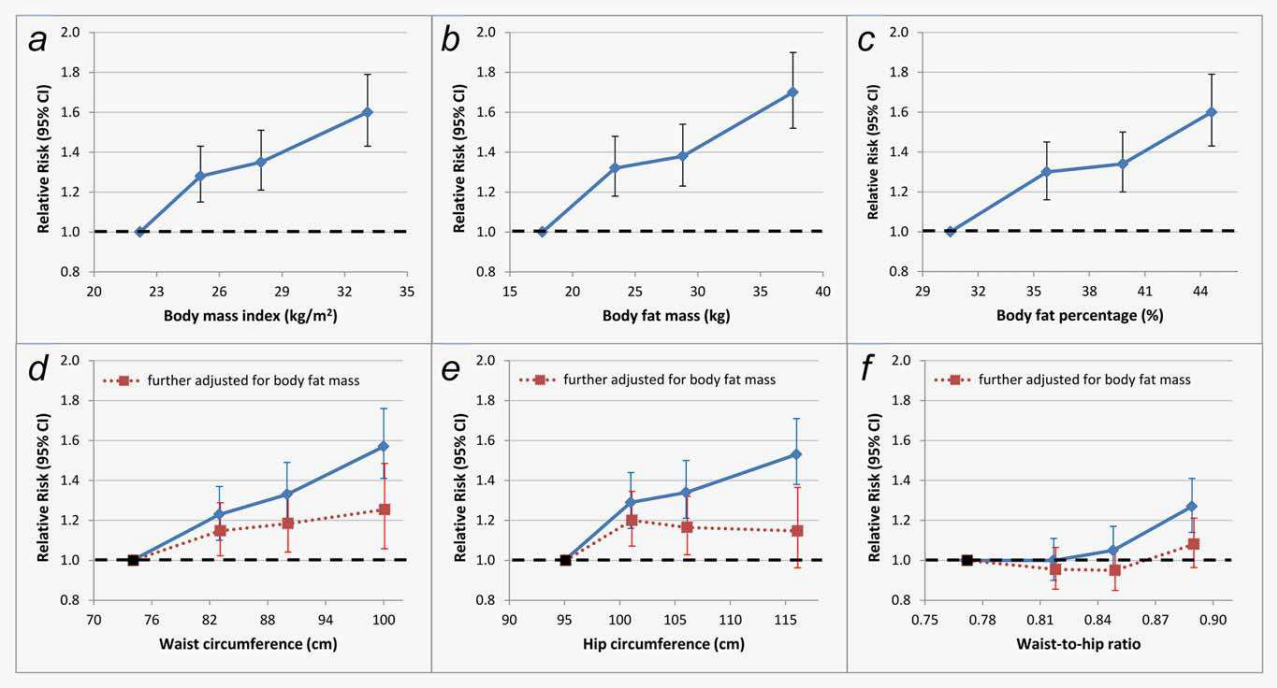


Figure 1. Association between various measures of adiposity and postmenopausal breast cancer risk: (*a*) body mass index, (*b*) body fat mass, (*c*) body fat percentage, (*d*) waist circumference, (*e*) hip circumference, (*f*) waist-to-hip ratio. Analyses are adjusted for family history of breast cancer, age at menarche, age at first birth, parity, age at menopause, previous HRT use, smoking, alcohol intake frequency, physical activity, ethnicity, height and body size at age 10. For panels *d*, *e* and *f*, the blue diamonds connected by a solid lines represent multivariable adjusted estimates and the red squares connected by dotted lines represent estimates further adjusted for body fat mass. Body size/composition measures were grouped into quartiles. The figure shows point estimates and 95% confidence intervals. [Color figure can be viewed at wileyonlinelibrary.com]

The magnitudes of risk associated with all measures of adiposity were greater in women who had been postmenopausal for 12 or more years; *p*-values for interaction were <0.02 for all adiposity variables except waist-to-hip ratio, BMI and body and trunk fat percentage (Table 5). For instance, comparing the top to the bottom quartile of body fat mass, postmenopausal women who were at least 12 years since menopause had a greater increased risk of breast cancer (RR, 1.86; 95% CI, 1.56–2.21) compared to those who were <12 years since menopause (RR, 1.47; 95% CI 1.24–1.75).

Discussion

In this large prospective study of postmenopausal women, we found statistically significant positive associations between all measures of adiposity and breast cancer risk. The relative risks, comparing the top with the bottom quartiles, were marginally greater for fat mass measured by impedance, but the RRs per SD increase and χ^2 statistics associated with each index of adiposity were of similar magnitude. All measures of central adiposity were also associated with postmenopausal breast cancer risk but these associations were attenuated after adjusting for body fat mass. Associations across all measures of adiposity were stronger in women who had been postmenopausal for ≥12 years.

Previous studies using impedance measures (based on 12,159; 7,523 and 13,598 participants) demonstrated a moderately greater magnitude of risk associated with postmenopausal breast cancer when comparing body fat to its surrogate measure, BMI.[20–22] Another study using DXA measures with 503 incident breast cancer cases concluded that there is no difference in the ability to predict postmenopausal breast cancer risk when comparing anthropometric indices to DXA-derived measures of body fat.[23] Our results are consistent with the findings from the latter study in that BMI and waist circumference were as informative for breast cancer risk as impedance measures of adiposity.

In these data, all measures of adiposity showed stronger associations with postmenopausal breast cancer risk among women who were ≥12 years since menopause compared to women with <12 years since menopause. These findings are consistent with those from two previous large prospective studies which found that the positive association between adiposity and postmenopausal breast cancer risk was more marked in women aged 65 or older[6] and in women who had experienced menopause ≥15 years previously.[22]

Elevated estrogen levels may stimulate carcinogenesis through increased cell proliferation as well as through pro-angiogenic and anti-apoptotic effects.[3,24] The relationship

Table 3. Association of height and overall adiposity measures with invasive breast cancer risk among never and former postmenopausal hormone replacement therapy users

					Stratified only	Multivariable-adjusted	Further adjusted for body size at age 10	Further adjusted for body size at age 10
Height quartiles	**Range (cm)**	**Median**	**Repeat median**	**Cases *n***	**RR (95% CI)[1]**	**RR (95% CI)[2]**	**RR (95% CI)[4]**	**RR (95% CI)[4]**
1	121–158	155	155	800	1	1	1	**Per SD increase**
2	158.1–162	161	160	720	1.07 (0.97-1.18)	1.07 (0.96-1.18)	1.07 (0.97-1.18)	1.10 (1.05-1.15)
3	162.1–166	164	164	653	1.09 (0.98-1.20)	1.08 (0.97-1.20)	1.08 (0.97-1.20)	
4	166.1–193	169	169	740	1.30 (1.18-1.44)	1.28 (1.16-1.42)	1.27 (1.15-1.41)	
χ^2								175.7
Weight quartiles	**Range (kg)**	**Median**	**Repeat median**	**Cases *n***	**RR (95% CI)[1]**	**RR (95% CI)[3]**	**RR (95% CI)[4]**	**RR (95% CI)[4]**
1	34.0–61.9	57.4	57.1	565	1	1	1	**Per SD increase**
2	62.0–69.2	65.6	65.4	715	1.27 (1.14-1.42)	1.25 (1.12-1.40)	1.27 (1.13-1.42)	1.24 (1.18-1.30)
3	69.3–78.5	73.4	72.6	738	1.31 (1.18-1.47)	1.29 (1.15-1.44)	1.32 (1.18-1.48)	
4	78.6–179.0	86.9	85.6	895	1.61 (1.45-1.79)	1.58 (1.41-1.76)	1.66 (1.48-1.86)	
χ^2								254.9
BMI quartiles	**Range (kg/m^2)**	**Median**	**Repeat median**	**Cases *n***	**RR (95% CI)[1]**	**RR (95% CI)[3]**	**RR (95% CI)[4]**	**RR (95% CI)[4]**
1	14.6–23.7	22.2	22.2	579	1	1	1	**Per SD increase**
2	23.8–26.4	25.1	25.1	743	1.25 (1.12-1.39)	1.26 (1.13-1.41)	1.28 (1.15-1.43)	1.21 (1.15-1.27)
3	26.5–29.9	28	28	753	1.29 (1.15-1.43)	1.32 (1.18-1.47)	1.35 (1.21-1.51)	
4	30.0–68.1	33.1	33	838	1.45 (1.30-1.62)	1.52 (1.36-1.70)	1.60 (1.43-1.79)	
χ^2								242.6
Body fat mass quartiles	**Range (kg)**	**Median**	**Repeat median**	**Cases *n***	**RR (95% CI)[1]**	**RR (95% CI)[3]**	**RR (95% CI)[4]**	**RR (95% CI)[4]**
1	8.0–20.5	17.3	17.6	537	1	1	1	**Per SD increase**
2	20.6–25.8	23.2	23.4	727	1.33 (1.19-1.49)	1.31 (1.17-1.47)	1.32 (1.18-1.48)	1.25 (1.19-1.31)
3	25.9–32.5	28.8	28.8	756	1.37 (1.23-1.53)	1.35 (1.21-1.51)	1.38 (1.23-1.54)	
4	32.6–108.4	38.5	37.6	893	1.64 (1.48-1.83)	1.62 (1.45-1.82)	1.70 (1.52-1.90)	
χ^2								260.3
Body fat % quartiles	**Range (%)**	**Median**	**Repeat median**	**Cases *n***	**RR (95% CI)[1]**	**RR (95% CI)[3]**	**RR (95% CI)[4]**	**RR (95% CI)[4]**
1	9.4–32.9	29.7	30.5	550	1	1	1	**Per SD increase**
2	33.0–37.4	35.4	35.7	729	1.30 (1.16-1.45)	1.29 (1.15-1.44)	1.30 (1.16-1.45)	1.23 (1.17-1.29)
3	37.5–41.8	39.5	39.8	763	1.33 (1.19-1.48)	1.32 (1.18-1.48)	1.34 (1.20-1.50)	
4	41.9–69.8	44.8	44.6	871	1.54 (1.38-1.72)	1.55 (1.39-1.73)	1.60 (1.43-1.79)	
χ^2								249.4

[1]Stratified by age at recruitment, region of recruitment and socioeconomic status (Townsend deprivation index).
[2]Adjusted for family history of breast cancer, age at menarche, age at first birth, parity, age at menopause, previous HRT use, smoking, alcohol intake frequency, physical activity and ethnicity.
[3]Adjusted for family history of breast cancer, age at menarche, age at first birth, parity, age at menopause, previous HRT use, smoking, alcohol intake frequency, physical activity ethnicity and height.
[4]Further adjusted for body size at age 10.

of adiposity with breast cancer risk in postmenopausal women can largely be explained by increased endogenous estrogen and decreased sex hormone-binding globulin. In two large studies with prospectively measured estrogen, adjustment for free estradiol completely attenuated the association between adiposity and breast cancer risk, suggesting that this relationship may be mediated by the concentration of bioavailable estradiol.[25,26] Our results further

Table 4. Association of central adiposity measures and invasive breast cancer risk among never and former postmenopausal HRT users

					Stratified only	Multivariable-adjusted	Further adjusted for body size at age 10	Further adjusted for body size at age 10	Further adjusted for body fat mass
Waist quartiles	**Range (cm)**	**Median**	**Repeat median**	**Cases *n***	**RR (95% CI)[1]**	**RR (95% CI)[2]**	**RR (95% CI)[3]**	**RR (95% CI)[3]**	**RR (95% CI)[4]**
1	48–76	72	74	576	1	1	1	**Per SD increase**	**Per SD increase**
2	76.1–84	81	83	757	1.23 (1.10–1.37)	1.22 (1.09–1.36)	1.23 (1.10–1.37)	1.23 (1.17–1.29)	1.07 (0.99–1.16)
3	84.1–93	89	90	750	1.33 (1.19–1.48)	1.31 (1.18–1.47)	1.33 (1.19–1.49)		
4	93.1–171	101	100	830	1.52 (1.37–1.69)	1.51 (1.36–1.69)	1.57 (1.41–1.76)		
χ^2								246.3	
Hip quartiles	**Range (cm)**	**Median**	**Repeat median**	**Cases *n***	**RR (95% CI)[1]**	**RR (95% CI)[2]**	**RR (95% CI)[3]**	**RR (95% CI)[3]**	**RR (95% CI)[4]**
1	62.0–97.0	94	95	658	1	1	1	**Per SD increase**	**Per SD increase**
2	97.1–102.0	100	101	693	1.31 (1.17–1.45)	1.28 (1.15–1.42)	1.29 (1.16–1.44)	1.21 (1.15–1.27)	1.01 (0.92–1.10)
3	102.1–109	105	106	772	1.35 (1.22–1.50)	1.32 (1.19–1.47)	1.34 (1.21–1.50)		
4	109.1–179	116	116	790	1.50 (1.35–1.66)	1.47 (1.32–1.63)	1.53 (1.38–1.71)		
χ^2								235.8	
WHR quartiles	**Range**	**Median**	**Repeat median**	**Cases *n***	**RR (95% CI)[1]**	**RR (95% CI)[2]**	**RR (95% CI)[3]**	**RR (95% CI)[3]**	**RR (95% CI)[4]**
1	0.503–0.773	0.745	0.771	673	1	1	1	**Per SD increase**	**Per SD increase**
2	0.775–0.820	0.798	0.817	680	0.99 (0.89–1.11)	1.0 (0.90–1.11)	1.00 (0.90–1.11)	1.14 (1.08–1.21)	1.03 (0.97–1.10)
3	0.820–0.870	0.843	0.848	718	1.03 (0.93–1.15)	1.04 (0.94–1.16)	1.05 (0.94–1.17)		
4	0.870–1.351	0.905	0.889	842	1.23 (1.11–1.37)	1.25 (1.13–1.39)	1.27 (1.14–1.41)		
χ^2								202.0	
Trunk fat mass quartiles	**Range (kg)**	**Median**	**Repeat median**	**Cases *n***	**RR (95% CI)[4]**	**RR (95% CI)[1]**	**RR (95% CI)[2]**	**RR (95% CI)[2]**	
1	0.6–10.2	8.4	8.6	545	1	1	1	**Per SD increase**	
2	10.3–13.3	11.8	12	731	1.29 (1.15-1.44)	1.27 (1.13-1.42)	1.28 (1.14-1.43)	1.25 (1.19-1.31)	
3	13.4–16.9	15	14.8	741	1.33 (1.19-1.48)	1.30 (1.16-1.46)	1.33 (1.18-1.49)		
4	17.0–48.3	19.9	19.3	896	1.66 (1.49-1.84)	1.62 (1.45-1.81)	1.69 (1.51-1.89)		
χ^2								259.4	
Trunk fat % quartiles	**Range (%)**	**Median**	**Repeat median**	**Cases *n***	**RR (95% CI)[4]**	**RR (95% CI)[1]**	**RR (95% CI)[2]**	**RR (95% CI)[2]**	
1	3.0–30.0	26.2	27.2	540	1	1	1	**Per SD increase**	
2	30.1–35.2	32.8	33.5	739	1.33 (1.19-1.48)	1.31 (1.17-1.47)	1.32 (1.18-1.47)	1.25 (1.18-1.31)	
3	35.3–40.0	37.5	37.6	753	1.36 (1.22-1.52)	1.34 (1.20-1.50)	1.36 (1.21-1.52)		
4	40.1–75.6	43.3	42.5	881	1.61 (1.45-1.80)	1.59 (1.42-1.77)	1.63 (1.46-1.83)		
χ^2								250.9	

[1]Stratified by age at recruitment, region of recruitment and socioeconomic status (Townsend deprivation index).
[2]Adjusted for family history of breast cancer, age at menarche, age at first birth, parity, age at menopause, previous HRT use, smoking, alcohol intake frequency, physical activity, height and ethnicity.
[3]Further adjusted for body size at age 10.
[4]Further adjusted for body fat mass.

suggest that the excess risk associated with increased levels of endogenous estrogens may take several years to fully outweigh the reduction in risk associated with adiposity in premenopausal women.[13,27]

A 2003 review on central obesity found that waist circumference and waist-to-hip ratio were no longer associated with postmenopausal breast cancer risk after adjusting for BMI or weight,[28] and some subsequent studies showed similar

Table 5. Association of body size/composition indices with invasive breast cancer risk in women who have been postmenopausal for <12 years *vs*. 12 or more years

	<12 years since menopause			≥12 years since menopause		
	Quartile	Cases, *n*	HR (95% CI)	Quartile	Cases, *n*	HR (95% CI)
Weight	1	288	1	1	283	1
	2	324	1.18 (0.99–1.40)	2	331	1.31 (1.10–1.55)
	3	284	1.05 (0.88–1.26)	3	387	1.56 (1.32–1.85)
	4	339	1.35 (1.14–1.61)	4	386	1.80 (1.52–2.13)
			p for interaction = 0.0049			
Body mass index	1	332	1	1	273	1
	2	317	1.10 (0.93–1.30)	2	357	1.40 (1.18–1.66)
	3	298	1.18 (1.00–1.41)	3	373	1.49 (1.26–1.77)
	4	298	1.28 (1.08–1.53)	4	383	1.83 (1.54–2.17)
			p for interaction = 0.0702			
Body fat mass	1	292	1	1	255	1
	2	326	1.25 (1.06–1.49)	2	336	1.39 (1.16–1.66)
	3	266	1.08 (0.90–1.29)	3	388	1.66 (1.40–1.97)
	4	337	1.47 (1.24–1.75)	4	383	1.86 (1.56–2.21)
			p for interaction = 0.0009			
Body fat percentage	1	309	1	1	255	1
	2	315	1.19 (1.01–1.41)	2	336	1.36 (1.14–1.62)
	3	290	1.18 (0.99–1.40)	3	373	1.50 (1.26–1.79)
	4	307	1.36 (1.15–1.62)	4	398	1.79 (1.51–2.13)
			p for interaction = 0.1106			
Waist	1	322	1	1	251	1
	2	340	1.16 (0.99–1.37)	2	373	1.37 (1.15–1.63)
	3	278	1.13 (0.95–1.35)	3	390	1.64 (1.38–1.94)
	4	296	1.32 (1.11–1.57)	4	374	1.83 (1.54–2.18)
			p for interaction = 0.0108			
Hip	1	336	1	1	312	1
	2	299	1.17 (0.99–1.38)	2	342	1.49 (1.26–1.76)
	3	305	1.13 (0.96–1.34)	3	385	1.62 (1.37–1.90)
	4	296	1.29 (1.09–1.53)	4	349	1.73 (1.46–2.04)
			p for interaction = 0.0167			
Waist-to-hip ratio	1	346	1	1	300	1
	2	314	1.02 (0.86–1.20)	2	325	1.04 (0.88–1.23)
	3	277	0.97 (0.81–1.14)	3	350	1.12 (0.95–1.33)
	4	299	1.17 (0.99–1.38)	4	413	1.36 (1.16–1.60)
			p for interaction = 0.4025			
Trunk fat mass	1	295	1	1	254	1
	2	319	1.20 (1.01–1.43)	2	348	1.37 (1.15–1.64)
	3	269	1.08 (0.90–1.29)	3	379	1.61 (1.35–1.91)
	4	336	1.43 (1.21–1.70)	4	381	1.87 (1.57–2.23)
			p for interaction = 0.0094			
Trunk fat percentage	1	294	1	1	255	1
	2	320	1.25 (1.05–1.48)	2	346	1.37 (1.14–1.63)
	3	293	1.22 (1.02–1.45)	3	361	1.51 (1.27–1.80)
	4	313	1.42 (1.19–1.69)	4	400	1.82 (1.53–2.16)
			p for interaction = 0.1580			

All analyses adjusted for family history of breast cancer, age at menarche, age at first birth, parity, age at menopause, previous HRT use, smoking, alcohol intake frequency, physical activity, height, ethnicity and body size at age 10.

results.[13,29,30] However, a recent meta-analysis found that the association between waist circumference and postmenopausal breast cancer risk was only slightly attenuated after adjusting for BMI.[9] In our analyses, which were adjusted for measured body fat mass rather than BMI, all associations between postmenopausal breast cancer and anthropometric measures of central adiposity were attenuated after adjustment for overall body fat, suggesting that overall body fat is the underlying adiposity-related predictor of postmenopausal breast cancer risk.

Recent Mendelian randomization studies report an inverse association between BMI predicted using single nucleotide polymorphisms identified by genome-wide association studies and postmenopausal breast cancer risk.[31,32] However, the BMI genetic score used in these studies may be more predictive of BMI in early-life rather than later adulthood.[33] In agreement with prior studies,[17,34] we found a protective association between greater childhood adiposity and breast cancer risk. We subsequently adjusted for comparative body size at age 10 as an independent risk factor for breast cancer and present relative risks both before and after this adjustment.

A notable strength of our study is the availability of objectively measured, rather than self-reported, adult body size and composition measures. Self-reported measures of adiposity such as weight tend to be underestimated, especially among overweight and obese women.[10] Other strengths include our prospective design, large sample size and virtually complete follow-up.[35] The availability of a wide range of known and putative risk factors for breast cancer allowed us to address potential bias and confounding.

Bioelectrical impedance measurements have been validated against hydrodensitometry, which is regarded as the gold standard for measuring body fat.[36,37] Although bioelectrical impedance measurements capture body fat more accurately than anthropometric measurements,[8,38] they have limitations. Studies assessing bioelectrical impedance in diverse populations may require population-specific calibration equations for different ethnic groups.[8] This is of limited concern in our study since over 95% of participants are White and we have adjusted for ethnic group in our analyses. Hydration status can also affect impedance analysis.[39] However, the accuracy of body composition measurements was improved by using a standardized measurement protocol administered by trained staff as well as utilizing repeat assessment data on all the body size and composition measures. Another limitation of the study is the lack of information on hormone receptor status of the tumor; however, since 70–80% of breast cancers included in our study of postmenopausal women are likely to be hormone receptor positive, the results are largely representative of hormonally responsive breast cancers.[40]

In conclusion, in this large prospective study, we found strong positive associations between all body size/composition variables and postmenopausal breast cancer risk. Anthropometric indices, with the exception of waist-to-hip ratio, were as informative for breast cancer risk as impedance measures of fat mass. Our results also suggest no association between postmenopausal breast cancer risk and abdominal adiposity beyond its contribution to overall adiposity.

References

1. Amadou A, Hainaut P, Romieu I. Role of obesity in the risk of breast cancer: lessons from anthropometry. *J Oncol* 2013;2013. Article ID 90649.
2. Grodin J, Siiteri P, MacDonald P. Source of estrogen production in postmenopausal women. *J Clin Endocrinol Metab* 1973;36:207–14.
3. Pike MC, Spicer DV, Dahmoush L, et al. Estrogens, progestogens, normal breast cell proliferation, and breast cancer risk. *Epidemiol Rev* 1993; 15:17–35.
4. Morimoto LM, White E, Chen Z, et al. Obesity, body size, and risk of postmenopausal breast cancer: the Women's Health Initiative. *Cancer Causes Control* 2002;13:741–51.
5. Canchola AJ, Anton-Culver H, Bernstein L, et al. Body size and the risk of postmenopausal breast cancer subtypes in the California Teachers Study cohort. *Cancer Causes Control* 2012;23:473–85.
6. Ritte R, Lukanova A, Berrino F, et al. Adiposity, hormone replacement therapy use and breast cancer risk by age and hormone receptor status: a large prospective cohort study. *Breast Cancer Res* 2012;14:R76.
7. James FR, Wootton S, Jackson A, et al. Obesity in breast cancer – what is the risk factor? *Eur J Cancer* 2015;51:705–20.
8. Bohm A, Heitmann BL. The use of bioelectrical impedance analysis for body composition in epidemiological studies. *Eur J Clin Nutr* 2013;67: S79–S85.
9. Chen GC, Chen SJ, Zhang R, et al. Central obesity and risks of pre- and postmenopausal breast cancer: a dose–response meta-analysis of prospective studies. *Obes Rev* 2016;17:1167–77.
10. Connor Gorber S, Tremblay M, Moher D, et al. A comparison of direct vs. self-report measures for assessing height, weight and body mass index: a systematic review. *Obes Rev* 2007;8:307–26.
11. UK Biobank. Protocol for a large scale prospective epidemiological resource. 2006. http://www.ukbiobank.ac.uk/resources/.
12. Million Women Study Collaborators. Breast cancer and hormone-replacement therapy in the Million Women Study. *Lancet* 2003;362: 419–27.
13. Lahmann PH, Hoffmann K, Allen N, et al. Body size and breast cancer risk: findings from the European Prospective Investigation into Cancer and Nutrition (EPIC). *Int J Cancer* 2004;111:762–71.
14. Collaborative Group of Hormonal Factors in Breast Cancer. Breast cancer and hormone replacement therapy: collaborative reanalysis of data from 51 epidemiological studies of 52,705 women with breast cancer and 108,411 women without breast cancer. *Lancet* 1997;350:1047–59.
15. MacMahon S, Peto R, Collins R, 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. *Lancet* 1990;335: 765–74.
16. Morris R, Carstairs V. Which deprivation? A comparison of selected deprivation indexes. *J Public Health Med* 1991;13:318–26.
17. Baer HJ, Tworoger SS, Hankinson SE, et al. Body fatness at young ages and risk of breast cancer throughout life. *Am J Epidemiol* 2010;171:1183–94.
18. Parish S, Peto R, Palmer A, et al. The joint effects of apolipoprotein B, apolipoprotein A1, LDL cholesterol, and HDL cholesterol on risk: 3510 cases of acute myocardial infarction and 9805 controls. *Eur Heart J* 2009;30:2137–46.
19. Alberti KG, Eckel RH, Grundy SM, et al. Harmonizing the metabolic syndrome: a joint interim statement of the International Diabetes Federation Task Force on Epidemiology and Prevention; National Heart, Lung, and Blood Institute; American Heart Association; World Heart Federation; International Atherosclerosis Society; and International Association for the Study of Obesity. *Circulation* 2009;120:1640–5.
20. Lahmann PH, Lissner L, Gullberg B, et al. A prospective study of adiposity and postmenopausal

breast cancer risk: the Malmo Diet and Cancer Study. *Int J Cancer* 2003;103:246–52.
21. Krebs EE, Taylor BC, Cauley JA, et al. Measures of adiposity and risk of breast cancer in older postmenopausal women. *J Am Geriatr Soc* 2006;54:63–9.
22. Macinnis RJ, English DR, Gertig DM, et al. Body size and composition and risk of postmenopausal breast cancer. *Cancer Epidemiol Biomarkers Prev* 2004;13:2117–25.
23. Rohan TE, Heo M, Choi L, et al. Body fat and breast cancer risk in postmenopausal women: a longitudinal study. *J Cancer Epidemiol* 2013;2013: 754815.
24. Losordo DW, Isner JM. Estrogen and angiogenesis: a review. *Arterioscler Thromb Vasc Biol* 2001; 21:6–12.
25. Rinaldi S, Key TJ, Peeters PH, et al. Anthropometric measures, endogenous sex steroids and breast cancer risk in postmenopausal women: a study within the EPIC cohort. *Int J Cancer* 2006; 118:2832–9.
26. Endogenous Hormones and Breast Cancer Collaborative Group. Body mass index, serum sex hormones, and breast cancer risk in postmenopausal women. *J Natl Cancer Inst* 2003;95:1218–26.
27. Huang Z, Hankinson SE, Colditz GA, et al. Dual effects of weight and weight gain on breast cancer risk. *JAMA* 1997;278:1407–11.
28. Harvie M, Hooper L, Howell AH. Central obesity and breast cancer risk: a systematic review. *Obes Rev* 2003;4:157–73.
29. Gaudet MM, Carter BD, Patel AV, et al. Waist circumference, body mass index, and postmenopausal breast cancer incidence in the Cancer Prevention Study-II Nutrition Cohort. *Cancer Causes Control* 2014;25:737–45.
30. Sellers TA, Davis J, Cerhan JR, et al. Interaction of waist/hip ratio and family history on the risk of hormone receptor-defined breast cancer in a prospective study of postmenopausal women. *Am J Epidemiol* 2002;155:225–33.
31. Guo Y, Warren Andersen S, Shu XO, et al. Genetically predicted body mass index and breast cancer risk: Mendelian randomization analyses of data from 145,000 women of European descent. *PLoS Med* 2016;13:e1002105.
32. Gao C, Patel CJ, Michailidou K, et al. Mendelian randomization study of adiposity-related traits and risk of breast, ovarian, prostate, lung and colorectal cancer. *Int J Epidemiol* 2016;45: 896–908.
33. Rukh G, Ahmad S, Ericson U, et al. Inverse relationship between a genetic risk score of 31 BMI loci and weight change before and after reaching middle age. *Int J Obes Relat Metab Disord* 2016; 40:252–9.
34. Song M, Willett WC, Hu FB, et al. Trajectory of body shape across the lifespan and cancer risk. *Int J Cancer* 2016;138:2383–95.
35. Moller H, Richards S, Hanchett N, et al. Completeness of case ascertainment and survival time error in English cancer registries: impact on 1-year survival estimates. *Br J Cancer* 2011;105: 170–6.
36. Brodie D, Moscrip V, Hutcheon R. Body composition measurement: a review of hydrodensitometry, anthropometry, and impedance methods. *Nutrition* 1998;14:296–310.
37. Lukaski HC, Johnson PE, Bolonchuk WW, et al. Assessment of fat-free mass using bioelectrical impedance measurements of the human body. *Am J Clin Nutr* 1985;41:810–7.
38. Houtkooper LB, Lohman TG, Going SB, et al. Why bioelectrical impedance analysis should be used for estimating adiposity. *Am J Clin Nutr* 1996;64:436s–48s.
39. NIH Consensus Statement. Bioelectrical impedance analysis in body composition measurement. National Institutes of Health Technology Assessment Conference Statement. December 12–14, 1994, *Nutrition* 1996;12:749–62.
40. Keen JC, Davidson NE. The biology of breast carcinoma. *Cancer* 2003;97:825–33.

Bibliography

[1] Global Burden of Disease Cancer Collaboration. Global, regional, and national cancer incidence, mortality, years of life lost, years lived with disability, and disability-adjusted life-years for 32 cancer groups, 1990 to 2015: A systematic analysis for the global burden of disease study. *JAMA Oncol*, 3(4): 524–548, 2017.

[2] T. J. Key, P. K. Verkasalo, and E. Banks. Epidemiology of breast cancer. *Lancet Oncol*, 2(3):133–40, 2001.

[3] D. M. Parkin, L. Boyd, and L. C. Walker. 16. the fraction of cancer attributable to lifestyle and environmental factors in the UK in 2010. *Br J Cancer*, 105 Suppl 2:S77–81, 2011.

[4] F. R. James, S. Wootton, A. Jackson, M. Wiseman, E. R. Copson, and R. I. Cutress. Obesity in breast cancer – what is the risk factor? *European Journal of Cancer*, 51(6):705–720, 2015.

[5] A. Bohm and B. L. Heitmann. The use of bioelectrical impedance analysis for body composition in epidemiological studies. *Eur J Clin Nutr*, 67 Suppl 1:S79–85, 2013.

[6] G. C. Chen, S. J. Chen, R. Zhang, K. Hidayat, J. B. Qin, et al. Central obesity and risks of pre- and postmenopausal breast cancer: a dose-response meta-analysis of prospective studies. *Obes Rev*, 17(11):1167–1177, 2016.

[7] J. E. Donnelly, S. N. Blair, J. M. Jakicic, M. M. Manore, J. W. Rankin, and B. K. Smith. American College of Sports Medicine position stand. Appropriate physical activity intervention strategies for weight loss and prevention of weight regain for adults. *Med Sci Sports Exerc*, 41(2):459–71, 2009.

[8] V. A. Catenacci and H. R. Wyatt. The role of physical activity in producing and maintaining weight loss. *Nat Clin Pract End Met*, 3(7):518–529, 2007.

[9] B. E. Ainsworth, C. J. Caspersen, C. E. Matthews, L. C. Masse, T. Baranowski, and W. Zhu. Recommendations to improve the accuracy of estimates of physical activity derived from self report. *J Phys Act Health*, 9 Suppl 1: S76–84, 2012.

[10] B. Lynch and M. F. Leitzmann. An evaluation of the evidence relating to physical inactivity, sedentary behavior, and cancer incidence and mortality. *Current Epidemiology Reports*, 4(3):221–231, 2017.

[11] J. Grodin, P. Siiteri, and P. MacDonald. Source of estrogen production in postmenopausal women. *J Clin Endocrinol Metab*, 36(2):207–214, 1973.

[12] D. O. Okorodudu, M. F. Jumean, V. M. Montori, A. Romero-Corral, V. K. Somers, et al. Diagnostic performance of body mass index to identify obesity as defined by body adiposity: a systematic review and meta-analysis. *Int J Obes*, 34(5):791–9, 2010.

[13] I. S. Fentiman, A. Fourquet, and G. N. Hortobagyi. Male breast cancer. *Lancet*, 367(9510):595–604, 2006.

[14] L. A. Torre, F. Bray, R. L. Siegel, J. Ferlay, J. Lortet-Tieulent, and A. Jemal. Global cancer statistics, 2012. *CA Cancer J Clin*, 65(2):87–108, 2015.

[15] R. G. Dumitrescu and I. Cotarla. Understanding breast cancer risk - where do we stand in 2005? *J Cell Mol Med*, 9(1):208–221, 2005.

[16] F. Bray, P. McCarron, and D. M. Parkin. The changing global patterns of female breast cancer incidence and mortality. *Breast Cancer Res*, 6(6): 229–39, 2004.

[17] L. A. Torre, R. L. Siegel, E. M. Ward, and A. Jemal. Global cancer incidence and mortality rates and trends–an update. *Cancer Epidemiol Biomarkers Prev*, 25(1):16–27, 2016.

[18] B. Zhang, X. O. Shu, R. J. Delahanty, C. Zeng, K. Michailidou, et al. Height and breast cancer risk: Evidence from prospective studies and mendelian randomization. *J Natl Cancer Inst*, 107(11), 2015.

[19] The Endogenous Hormones and Breast Cancer Collaborative Group. Insulin-like growth factor 1 (IGF1), IGF binding protein 3 (IGFBP3), and breast cancer risk: pooled individual data analysis of 17 prospective studies. *Lancet Oncol*, 11(6):530–542, 2010.

[20] N. C. Onland-Moret, P. H. Peeters, C. H. van Gils, F. Clavel-Chapelon, T. Key, et al. Age at menarche in relation to adult height: the EPIC study. *Am J Epidemiol*, 162(7):623–32, 2005.

[21] G. K. Reeves, K. Pirie, V. Beral, J. Green, E. Spencer, and D. Bull. Cancer incidence and mortality in relation to body mass index in the Million Women Study: cohort study. *BMJ*, 335(7630):1134, 2007.

[22] Endogenous Hormones and Breast Cancer Collaborative Group. Body mass index, serum sex hormones, and breast cancer risk in postmenopausal women. *J Natl Cancer Inst*, 95(16):1218–26, 2003.

[23] T. J. Key, N. E. Allen, E. A. Spencer, and R. C. Travis. Nutrition and breast cancer. *Breast*, 12(6):412–416, 2003.

[24] L. M. Morimoto, E. White, Z. Chen, R. T. Chlebowski, J. Hays, et al. Obesity, body size, and risk of postmenopausal breast cancer: the Women's Health Initiative. *Cancer Causes Control*, 13(8):741–51, 2002.

[25] K. Kerlikowske, C. C. Gard, J. A. Tice, E. Ziv, S. R. Cummings, and D. L. Miglioretti. Risk factors that increase risk of estrogen receptor-positive and -negative breast cancer. *J Natl Cancer Inst*, 109(5), 2017.

[26] K. B. Michels, K. L. Terry, and W. C. Willett. Longitudinal study on the role of body size in premenopausal breast cancer. *Arch Intern Med*, 166(21): 2395–402, 2006.

[27] H. B. Nichols, M. J. Schoemaker, L. B. Wright, C. McGowan, M. N. Brook, et al. The Premenopausal Breast Cancer Collaboration: A pooling project of studies participating in the National Cancer Institute Cohort Consortium. *Cancer Epidemiol Biomarkers Prev*, 26(9):1360–1369, 2017.

[28] N. Potischman, C. A. Swanson, P. Siiteri, and R. N. Hoover. Reversal of

relation between body mass and endogenous estrogen concentrations with menopausal status. *J Natl Cancer Inst*, 88(11):756–8, 1996.

[29] J. Ahn, A. Schatzkin, Jr. Lacey, J. V., D. Albanes, R. Ballard-Barbash, et al. Adiposity, adult weight change, and postmenopausal breast cancer risk. *Archives of Internal Medicine*, 167(19):2091–2102, 2007.

[30] A. J. Canchola, H. Anton-Culver, L. Bernstein, C. A. Clarke, K. Henderson, et al. Body size and the risk of postmenopausal breast cancer subtypes in the California Teachers Study cohort. *Cancer Causes Control*, 2012.

[31] R. Ritte, A. Lukanova, F. Berrino, L. Dossus, A. Tjonneland, et al. Adiposity, hormone replacement therapy use and breast cancer risk by age and hormone receptor status: a large prospective cohort study. *Breast Cancer Research*, 14(3), 2012.

[32] A. G. Renehan, M. Zwahlen, and M. Egger. Adiposity and cancer risk: new mechanistic insights from epidemiology. *Nat Rev Cancer*, 15(8):484–98, 2015.

[33] H. R. Harris, W. C. Willett, K. L. Terry, and K. B. Michels. Body fat distribution and risk of premenopausal breast cancer in the Nurses' Health Study ii. *J Natl Cancer Inst*, 103(3):273–8, 2011.

[34] H. J. Baer, S. S. Tworoger, S. E. Hankinson, and W. C. Willett. Body fatness at young ages and risk of breast cancer throughout life. *Am J Epidemiol*, 171(11):1183–94, 2010.

[35] Z. J. Andersen, J. L. Baker, K. Bihrmann, I. Vejborg, T. I. Sorensen, and E. Lynge. Birth weight, childhood body mass index, and height in relation

to mammographic density and breast cancer: a register-based cohort study. *Breast Cancer Res*, 16(1):R4, 2014.

[36] N. Mauras, R. J. Santen, G. Colon-Otero, J. Hossain, Q. Wang, et al. Estrogens and their genotoxic metabolites are increased in obese prepubertal girls. *J Clin Endocrinol Metab*, 100(6):2322–8, 2015.

[37] A. Cabanes, M. Wang, S. Olivo, S. DeAssis, J. A. Gustafsson, et al. Prepubertal estradiol and genistein exposures up-regulate BRCA1 mRNA and reduce mammary tumorigenesis. *Carcinogenesis*, 25(5):741–8, 2004.

[38] Y. Guo, S. Warren Andersen, X. O. Shu, K. Michailidou, M. K. Bolla, et al. Genetically predicted body mass index and breast cancer risk: Mendelian randomization analyses of data from 145,000 women of European descent. *PLoS Med*, 13(8):e1002105, 2016.

[39] C. Gao, C. J. Patel, K. Michailidou, U. Peters, J. Gong, et al. Mendelian randomization study of adiposity-related traits and risk of breast, ovarian, prostate, lung and colorectal cancer. *Int J Epidemiol*, 45(3):896–908, 2016.

[40] G. Rukh, S. Ahmad, U. Ericson, G. Hindy, T. Stocks, et al. Inverse relationship between a genetic risk score of 31 bmi loci and weight change before and after reaching middle age. *Int J Obes (Lond)*, 40(2):252–9, 2016.

[41] L. M. Welti, D. P. Beavers, B. J. Caan, H. Sangi-Haghpeykar, M. Z. Vitolins, and K. M. Beavers. Weight fluctuation and cancer risk in postmenopausal women: The Women's Health Initiative. *Cancer Epidemiol Biomarkers Prev*, 26(5):779–786, 2017.

[42] N. Keum, D. C. Greenwood, D. H. Lee, R. Kim, D. Aune, et al. Adult weight gain and adiposity-related cancers: a dose-response meta-analysis of prospective observational studies. *J Natl Cancer Inst*, 107(2), 2015.

[43] B. Rosner, A. H. Eliassen, A. T. Toriola, W. Y. Chen, S. E. Hankinson, et al. Weight and weight changes in early adulthood and later breast cancer risk. *Int J Cancer*, 140(9):2003–2014, 2017.

[44] H. K. Neilson, M. S. Farris, C. R. Stone, M. M. Vaska, D. R. Brenner, and C. M. Friedenreich. Moderate-vigorous recreational physical activity and breast cancer risk, stratified by menopause status: a systematic review and meta-analysis. *Menopause*, 24(3):322–344, 2017.

[45] P. K. Verkasalo, H. V. Thomas, P. N. Appleby, G. K. Davey, and T. J. Key. Circulating levels of sex hormones and their relation to risk factors for breast cancer: a cross-sectional study in 1092 pre- and postmenopausal women (United Kingdom). *Cancer Causes Control*, 12(1):47–59, 2001.

[46] C. H. van Gils, P. H. Peeters, M. C. Schoenmakers, R. M. Nijmeijer, N. C. Onland-Moret, et al. Physical activity and endogenous sex hormone levels in postmenopausal women: a cross-sectional study in the Prospect-EPIC Cohort. *Cancer Epidemiol Biomarkers Prev*, 18(2):377–83, 2009.

[47] C. M. Friedenreich, H. K. Neilson, C. G. Woolcott, Q. Wang, F. Z. Stanczyk, et al. Inflammatory marker changes in a yearlong randomized exercise intervention trial among postmenopausal women. *Cancer Prev Res (Phila)*, 5 (1):98–108, 2012.

[48] H. K. Neilson, S. M. Conroy, and C. M. Friedenreich. The influence of energetic factors on biomarkers of postmenopausal breast cancer risk. *Curr Nutr Rep*, 3:22–34, 2014.

[49] Food, nutrition, physical activity, and the prevention of cancer: a global perspective. Technical report, World Cancer Research Fund and American Institute for Cancer Research, 2007.

[50] M. J. Emaus, P. H. Peeters, M. F. Bakker, K. Overvad, A. Tjonneland, et al. Vegetable and fruit consumption and the risk of hormone receptor-defined breast cancer in the EPIC cohort. *Am J Clin Nutr*, 103(1):168–77, 2016.

[51] M. S. Farvid, W. Y. Chen, K. B. Michels, E. Cho, W. C. Willett, and H. Eliassen. Fruit and vegetable consumption in adolescence and early adulthood and risk of breast cancer: population based cohort study. *BMJ*, 353, 2016.

[52] J. Y. Dong, K. He, P. Wang, and L. Q. Qin. Dietary fiber intake and risk of breast cancer: a meta-analysis of prospective cohort studies. *Am J Clin Nutr*, 94(3):900–5, 2011.

[53] D. Aune, D. S. Chan, D. C. Greenwood, A. R. Vieira, D. A. Rosenblatt, et al. Dietary fiber and breast cancer risk: a systematic review and meta-analysis of prospective studies. *Ann Oncol*, 23(6):1394–402, 2012.

[54] K. E. Bradbury, P. N. Appleby, and T. J. Key. Fruit, vegetable, and fiber intake in relation to cancer risk: findings from the European Prospective Investigation into Cancer and Nutrition (EPIC). *Am J Clin Nutr*, 100 Suppl 1:394S–8S, 2014.

[55] A. Diallo, M. Deschasaux, P. Latino-Martel, S. Hercberg, P. Galan, et al. Red and processed meat intake and cancer risk: Results from the prospective NutriNet-Sante cohort study. *Int J Cancer*, 142(2):230–237, 2018.

[56] J. J. Anderson, N. D. M. Darwis, D. F. Mackay, C. A. Celis-Morales, D. M. Lyall, et al. Red and processed meat consumption and breast cancer: UK Biobank cohort study and meta-analysis. *Eur J Cancer*, 90:73–82, 2018.

[57] Y. Cao, L. Hou, and W. Wang. Dietary total fat and fatty acids intake, serum fatty acids and risk of breast cancer: A meta-analysis of prospective cohort studies. *Int J Cancer*, 138(8):1894–904, 2016.

[58] S. Jung, M. Wang, K. Anderson, L. Baglietto, L. Bergkvist, et al. Alcohol consumption and breast cancer risk by estrogen receptor status: in a pooled analysis of 20 studies. *Int J Epidemiol*, 45(3):916–28, 2016.

[59] R. Suzuki, N. Orsini, L. Mignone, S. Saji, and A. Wolk. Alcohol intake and risk of breast cancer defined by estrogen and progesterone receptor status–a meta-analysis of epidemiological studies. *Int J Cancer*, 122(8):1832–41, 2008.

[60] Endogenous Hormones and Breast Cancer Collaborative Group. Circulating sex hormones and breast cancer risk factors in postmenopausal women: reanalysis of 13 studies. *Br J Cancer*, 105(5):709–22, 2011.

[61] R. Baan, K. Straif, Y. Grosse, B. Secretan, F. El Ghissassi, et al. Carcinogenicity of alcoholic beverages. *Lancet Oncol*, 8(4):292–293, 2007.

[62] H. K. Seitz and F. Stickel. Molecular mechanisms of alcohol-mediated carcinogenesis. *Nat Rev Cancer*, 7(8):599–612, 2007.

[63] International Agency for Research on Cancer. A review of human carcinogens: Personal habits and indoor combustions. 100(Pt E):1–538, 2012.

[64] F. Xue, W. C. Willett, B. A. Rosner, S. E. Hankinson, and K. B. Michels. Cigarette smoking and the incidence of breast cancer. *Archives of Internal Medicine*, 171(2):125–133, 2011.

[65] M. M. Gaudet, B. D. Carter, L. A. Brinton, R. T. Falk, I. T. Gram, et al. Pooled analysis of active cigarette smoking and invasive breast cancer risk in 14 cohort studies. *Int J Epidemiol*, 46(3):881–893, 2017.

[66] M. M. Gaudet, S. M. Gapstur, J. Sun, W. R. Diver, L. M. Hannan, and M. J. Thun. Active smoking and breast cancer risk: Original cohort data and meta-analysis. *J Natl Cancer Inst*, 105(8):515–525, 2013.

[67] Collaborative Group on Hormonal Factors in Breast Cancer. Menarche, menopause, and breast cancer risk: individual participant meta-analysis, including 118,964 women with breast cancer from 117 epidemiological studies. *Lancet Oncol*, 13(11):1141–51, 2012.

[68] F. R. Day, D. J. Thompson, H. Helgason, D. I. Chasman, H. Finucane, et al. Genomic analyses identify hundreds of variants associated with age at menarche and support a role for puberty timing in cancer risk. *Nat Genet*, 49(6):834–841, 2017.

[69] C. Koprowski, R. K. Ross, W. J. Mack, B. E. Henderson, and L. Bernstein.

Diet, body size and menarche in a multiethnic cohort. *Br J Cancer*, 79 (11-12):1907–11, 1999.

[70] M. C. Pike, M. D. Krailo, B. E. Henderson, J. T. Casagrande, and D. G. Hoel. Hormonal risk factors, breast tissue age and the age-incidence of breast cancer. *Nature*, 303(5920):767–770, 1983.

[71] M. Garland, D. J. Hunter, G. A. Colditz, J. E. Manson, M. J. Stampfer, et al. Menstrual cycle characteristics and history of ovulatory infertility in relation to breast cancer risk in a large cohort of us women. *Am J Epidemiol*, 147(7):636–643, 1998.

[72] M. Lambe, C. Hsieh, D. Trichopoulos, A. Ekbom, M. Pavia, and H. O. Adami. Transient increase in the risk of breast cancer after giving birth. *N Engl J Med*, 331(1):5–9, 1994.

[73] J. L. Kelsey, M. D. Gammon, and E. M. John. Reproductive factors and breast cancer. *Epidemiol Rev*, 15(1):36–47, 1993.

[74] I. H. Russo and J. Russo. Pregnancy-induced changes in breast cancer risk. *Journal of Mammary Gland Biology and Neoplasia*, 16(3):221–233, 2011.

[75] L. Sivaraman and D. Medina. Hormone-induced protection against breast cancer. *Journal of Mammary Gland Biology and Neoplasia*, 7(1):77–92, 2002.

[76] C. M. Blakely, A. J. Stoddard, G. K. Belka, K. D. Dugan, K. L. Notarfrancesco, et al. Hormone-induced protection against mammary tumorigenesis is conserved in multiple rat strains and identifies a core gene expression signature induced by pregnancy. *Cancer Research*, 66(12):6421–6431, 2006.

[77] J. Russo, G. A. Balogh, and I. H. Russo. Full-term pregnancy induces a specific genomic signature in the human breast. *Cancer Epidemiology Biomarkers & Prevention*, 17(1):51–66, 2008.

[78] Collaborative Group on Hormonal Factors in Breast Cancer. Breast cancer and breastfeeding: 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*, 360(9328):187–195, 2002.

[79] Collaborative Group on Hormonal Factors in Breast. Breast cancer and hormonal contraceptives: collaborative reanalysis of individual data on 53,297 women with breast cancer and 100,239 women without breast cancer from 54 epidemiological studies. *Lancet*, 347(9017):1713–27, 1996.

[80] L. S. Morch, C. W. Skovlund, P. C. Hannaford, L. Iversen, S. Fielding, and O. Lidegaard. Contemporary hormonal contraception and the risk of breast cancer. *N Engl J Med*, 377(23):2228–2239, 2017.

[81] F. R. Day, K. S. Ruth, D. J. Thompson, K. L. Lunetta, N. Pervjakova, et al. Large-scale genomic analyses link reproductive aging to hypothalamic signaling, breast cancer susceptibility and BRCA1-mediated DNA repair. *Nat Genet*, 47(11):1294–1303, 2015.

[82] Collaborative Group on Hormonal Factors in Breast. Breast cancer and hormone replacement therapy: collaborative reanalysis of data from 51 epidemiological studies of 52,705 women with breast cancer and 108,411 women without breast cancer. *The Lancet*, 350(9084):1047–1059, 1997.

[83] R. T. Chlebowski, G. L. Anderson, M. Gass, D. S. Lane, A. K. Aragaki, et al. Estrogen plus progestin and breast cancer incidence and mortality in postmenopausal women. *JAMA*, 304(15):1684–92, 2010.

[84] Million Women Study Collaborators. Breast cancer and hormone-replacement therapy in the Million Women Study. *The Lancet*, 362(9382): 419–427, 2003.

[85] G. L. Anderson, M. Limacher, A. R. Assaf, T. Bassford, S. A. Beresford, et al. Effects of conjugated equine estrogen in postmenopausal women with hysterectomy: the Women's Health Initiative randomized controlled trial. *JAMA*, 291(14):1701–12, 2004.

[86] Beral V., Banks E., and Reeves G. Effects of estrogen-only treatment in postmenopausal women. *JAMA*, 292(6):683, 2004.

[87] M. E. Jones, M. J. Schoemaker, L. Wright, E. McFadden, J. Griffin, et al. Menopausal hormone therapy and breast cancer: what is the true size of the increased risk? *Br J Cancer*, 115(5):607–15, 2016.

[88] H. V. Thomas, G. K. Reeves, and T. J. Key. Endogenous estrogen and postmenopausal breast cancer: a quantitative review. *Cancer Causes Control*, 8 (6):922–8, 1997.

[89] L. Baglietto, G. Severi, D. R. English, K. Krishnan, J. L. Hopper, et al. Circulating steroid hormone levels and risk of breast cancer for postmenopausal women. *Cancer Epidemiology Biomarkers & Prevention*, 19(2):492–502, 2010.

[90] Endogenous Hormones and Breast Cancer Collaborative Group. Sex hormones and risk of breast cancer in premenopausal women: a collaborative reanalysis of individual participant data from seven prospective studies. *Lancet Oncol*, 14(10):1009–1019, 2013.

[91] W. Ge, T. V. Clendenen, Y. Afanasyeva, K. L. Koenig, C. Agnoli, et al. Circulating anti-mullerian hormone and breast cancer risk: A study in ten prospective cohorts. *Int J Cancer*, 142(11):2215–2226, 2018.

[92] S. L. Broer, F. J. Broekmans, J. S. Laven, and B. C. Fauser. Anti-mullerian hormone: ovarian reserve testing and its potential clinical implications. *Hum Reprod Update*, 20(5):688–701, 2014.

[93] K. N. Anderson, R. B. Schwab, and M. E. Martinez. Reproductive risk factors and breast cancer subtypes: a review of the literature. *Breast Cancer Res Treat*, 144(1):1–10, 2014.

[94] W. Y. Chen and G. A. Colditz. Risk factors and hormone-receptor status: epidemiology, risk-prediction models and treatment implications for breast cancer. *Nat Clin Pract Oncol*, 4(7):415–23, 2007.

[95] N. F. Boyd, H. Guo, L. J. Martin, L. Sun, J. Stone, et al. Mammographic density and the risk and detection of breast cancer. *N Engl J Med*, 356(3): 227–36, 2007.

[96] V. A. McCormack and I. dos Santos Silva. Breast density and parenchymal patterns as markers of breast cancer risk: a meta-analysis. *Cancer Epidemiol Biomarkers Prev*, 15(6):1159–69, 2006.

[97] M. S. Rice, K. A. Bertrand, T. J. VanderWeele, B. A. Rosner, X. Liao, et al. Mammographic density and breast cancer risk: a mediation analysis. *Breast Cancer Res*, 18(1):94, 2016.

[98] J. A. Tice, E. S. O'Meara, D. L. Weaver, C. Vachon, R. Ballard-Barbash, and K. Kerlikowske. Benign breast disease, mammographic breast density, and the risk of breast cancer. *J Natl Cancer Inst*, 105(14):1043–9, 2013.

[99] C. M. Ronckers, C. A. Erdmann, and C. E. Land. Radiation and breast cancer: a review of current evidence. *Breast Cancer Res*, 7(1):21–32, 2005.

[100] Collaborative Group on Hormonal Factors in Breast Cancer. Familial breast cancer: collaborative reanalysis of individual data from 52 epidemiological studies including 58,209 women with breast cancer and 101,986 women without the disease. *Lancet*, 358(9291):1389–99, 2001.

[101] G. A. Colditz, K. A. Kaphingst, S. E. Hankinson, and B. Rosner. Family history and risk of breast cancer: Nurses' Health Study. *Breast Cancer Res Treat*, 133(3):1097–104, 2012.

[102] N. Mavaddat, A. C. Antoniou, D. F. Easton, and M. Garcia-Closas. Genetic susceptibility to breast cancer. *Molecular Oncology*, 4(3):174–191, 2010.

[103] M. Ghoussaini, P. D. Pharoah, and D. F. Easton. Inherited genetic susceptibility to breast cancer: the beginning of the end or the end of the beginning? *Am J Pathol*, 183(4):1038–51, 2013.

[104] A. C. Antoniou and D. F. Easton. Models of genetic susceptibility to breast cancer. *Oncogene*, 25(43):5898–905, 2006.

[105] A. Rudolph, J. Chang-Claude, and M. K. Schmidt. Gene-environment interaction and risk of breast cancer. *Br J Cancer*, 114(2):125–33, 2016.

[106] S. Nickels, T. Truong, R. Hein, K. Stevens, K. Buck, et al. Evidence of gene-environment interactions between common breast cancer susceptibility loci and established environmental risk factors. *PLoS Genetics*, 9(3), 2013.

[107] M. Barrdahl, A. Rudolph, J. L. Hopper, M. C. Southey, A. Broeks, et al. Gene-environment interactions involving functional variants: Results from the Breast Cancer Association Consortium. *Int J Cancer*, 141(9):1830–1840, 2017.

[108] L. E. Johns, M. E. Jones, M. J. Schoemaker, E. McFadden, A. Ashworth, and A. J. Swerdlow. Domestic light at night and breast cancer risk: a prospective analysis of 105,000 UK women in the Generations Study. *Br J Cancer*, 118 (4):600–606, 2018.

[109] X. Zhang, S. A. Smith-Warner, L. C. Collins, B. Rosner, W. C. Willett, and S. E. Hankinson. Use of aspirin, other nonsteroidal anti-inflammatory drugs, and acetaminophen and postmenopausal breast cancer incidence. *J Clin Oncol*, 30(28):3468–77, 2012.

[110] UK Biobank. UK Biobank: Protocol for a large-scale prospective epidemiological resource, 2010.

[111] A. Doherty, D. Jackson, N. Hammerla, T. Plotz, P. Olivier, et al. Large scale population assessment of physical activity using wrist worn accelerometers: the UK Biobank study. *PLoS One*, 12(2):e0169649, 2017.

[112] N. E. Allen, C. Sudlow, T. Peakman, and R. Collins. UK Biobank data: come and get it. *Sci Transl Med*, 6(224):224ed4, 2014.

[113] UK Biobank data management & sharing plan. `http://www.ukbiobank.ac.uk/wp-content/uploads/2013/10/ukbiobank-data-management.pdf`, . Accessed: 2017-03-12.

[114] A. Fry, T. J. Littlejohns, C. Sudlow, N. Doherty, L. Adamska, et al. Comparison of sociodemographic and health-related characteristics of UK Biobank participants with those of the general population. *Am J Epidemiol*, 186(9): 1026–1034, 2017.

[115] UK Biobank: Imaging modality DXA. `https://biobank.ctsu.ox.ac.uk/crystal/docs/DXA_explan_doc.pdf`, . Accessed: 2017-03-12.

[116] UK Biobank axiom array:content summary. `http://www.ukbiobank.ac.uk/wp-content/uploads/2014/04/UK-Biobank-Axiom-Array-Content-Summary-2014.pdf`, . Accessed: 2017-03-12.

[117] Data providers and dates of data availability. `https://biobank.ctsu.ox.ac.uk/crystal/exinfo.cgi?src=Data_providers_and_dates`, . Accessed: 2017-03-12.

[118] ICD-10 version 2010. `http://apps.who.int/classifications/icd10/browse/2010/en#/C50-C50`. Accessed: 2017-03-12.

[119] R. Morris and V. Carstairs. Which deprivation? a comparison of selected deprivation indexes. *J Public Health Med*, 13(4):318–26, 1991.

[120] C. Sudlow, J. Gallacher, N. Allen, V. Beral, P. Burton, et al. UK Biobank: an open access resource for identifying the causes of a wide range of complex diseases of middle and old age. *PLoS Med*, 12(3):e1001779, 2015.

[121] R. M. Tamimi, D. Spiegelman, S. A. Smith-Warner, M. Wang, M. Pazaris, et al. Population attributable risk of modifiable and nonmodifiable breast cancer risk factors in postmenopausal breast cancer. *Am J Epidemiol*, 184 (12):884–893, 2016.

[122] M. Kyrgiou, I. Kalliala, G. Markozannes, M. J. Gunter, E. Paraskevaidis, et al. Adiposity and cancer at major anatomical sites: umbrella review of the literature. *Bmj*, 356:j477, 2017.

[123] M. S. R. Shawon, M. Eriksson, and J. Li. Body size in early life and risk of breast cancer. *Breast Cancer Res*, 19, 2017.

[124] Surgeon General. The health consequences of smoking—50 years of progress: A report of the Surgeon General. In *US Department of Health and Human Services*, 2014.

[125] M. Ewertz, S. W. Duffy, H. O. Adami, G. Kvale, E. Lund, et al. Age at first birth, parity and risk of breast cancer: a meta-analysis of 8 studies from the Nordic countries. *Int J Cancer*, 46(4):597–603, 1990.

[126] Independent UK Panel on Breast Cancer Screening. The benefits and harms of breast cancer screening: an independent review. *Lancet*, 380(9855):1778–86, 2012.

[127] H. Moller, S. Richards, N. Hanchett, S. P. Riaz, M. Luchtenborg, et al.

Completeness of case ascertainment and survival time error in English cancer registries: impact on 1-year survival estimates. *Br J Cancer*, 105(1):170–6, 2011.

[128] A. Must, J. Spadano, E. H. Coakley, A. E. Field, G. Colditz, and W. H. Dietz. The disease burden associated with overweight and obesity. *JAMA*, 282(16):1523–1529, 1999.

[129] K. A. Britton, I. M. Lee, L. Wang, J. M. Gaziano, J. E. Manson, et al. Physical activity and the risk of becoming overweight or obese in middle-aged and older women. *Obesity (Silver Spring)*, 20(5):1096–103, 2012.

[130] U. Ekelund, H. Besson, J. Luan, A. M. May, S. J. Sharp, et al. Physical activity and gain in abdominal adiposity and body weight: prospective cohort study in 288,498 men and women. *Am J Clin Nutr*, 93(4):826–35, 2011.

[131] W. M. Kohrt, K. A. Obert, and J. O. Holloszy. Exercise training improves fat distribution patterns in 60- to 70-year-old men and women. *J Gerontol*, 47(4):M99–105, 1992.

[132] R. S. Schwartz, W. P. Shuman, V. Larson, K. C. Cain, G. W. Fellingham, et al. The effect of intensive endurance exercise training on body fat distribution in young and older men. *Metabolism*, 40(5):545–51, 1991.

[133] H. W. Peters, I. C. Westendorp, A. E. Hak, D. E. Grobbee, C. D. Stehouwer, and J. C. others. Menopausal status and risk factors for cardiovascular disease. *J Intern Med*, 246(6):521–8, 1999.

[134] K. R. Evenson, S. Wilcox, M. Pettinger, R. Brunner, A. C. King, and A. McTiernan. Vigorous leisure activity through women's adult life: the Women's Health Initiative observational cohort study. *Am J Epidemiol*, 156(10):945–53, 2002.

[135] C. L. Craig, A. L. Marshall, M. Sjostrom, A. E. Bauman, M. L. Booth, et al. International physical activity questionnaire: 12-country reliability and validity. *Medicine & Science in Sports & Exercise*, 35(8):1381–95, 2003.

[136] B. E. Ainsworth, W. L. Haskell, S. D. Herrmann, N. Meckes, Jr. Bassett, D. R., et al. 2011 compendium of physical activities: a second update of codes and MET values. *Medicine & Science in Sports & Exercise*, 43(8): 1575–81, 2011.

[137] M. E. Armstrong, B. J. Cairns, J. Green, G. K. Reeves, and V. Beral. Reported frequency of physical activity in a large epidemiological study: relationship to specific activities and repeatability over time. *BMC Med Res Methodol*, 11:97, 2011.

[138] P. Imbeault, S. Saint-Pierre, N. Almeras, and A. Tremblay. Acute effects of exercise on energy intake and feeding behaviour. *Br J Nutr*, 77(4):511–21, 1997.

[139] David S. Gray and Ken Fujioka. Use of relative weight and body mass index for the determination of adiposity. *Journal of Clinical Epidemiology*, 44(6): 545–550, 1991.

[140] M. Neovius, Y. Linne, and S. Rossner. BMI, waist-circumference and waist-

hip-ratio as diagnostic tests for fatness in adolescents. *Int J Obes (Lond)*, 29 (2):163–9, 2005.

[141] C. Ranasinghe, P. Gamage, P. Katulanda, N. Andraweera, S. Thilakarathne, and P. Tharanga. Relationship between body mass index (bmi) and body fat percentage, estimated by bioelectrical impedance, in a group of Sri Lankan adults: a cross sectional study. *BMC Public Health*, 13:797, 2013.

[142] A. Amadou, P. Hainaut, and I. Romieu. Role of obesity in the risk of breast cancer: lessons from anthropometry. *J Oncol*, 2013:906495, 2013.

[143] B. A. Irving, C. K. Davis, D. W. Brock, J. Y. Weltman, D. Swift, et al. Effect of exercise training intensity on abdominal visceral fat and body composition. *Medicine & Science in Sports & Exercise*, 40(11):1863–72, 2008.

[144] J. K. Cho, S. H. Lee, J. Y. Lee, and H. S. Kang. Randomized controlled trial of training intensity in adiposity. *Int J Sports Med*, 32(6):468–75, 2011.

[145] B. J. Nicklas, X. Wang, T. You, M. F. Lyles, J. Demons, et al. Effect of exercise intensity on abdominal fat loss during calorie restriction in overweight and obese postmenopausal women: a randomized, controlled trial. *Am J Clin Nutr*, 89(4):1043–52, 2009.

[146] C. A. Slentz, L. B. Aiken, J. A. Houmard, C. W. Bales, J. L. Johnson, et al. Inactivity, exercise, and visceral fat. strride: a randomized, controlled study of exercise intensity and amount. *J Appl Physiol (1985)*, 99(4):1613–8, 2005.

[147] H. O. Chambliss. Exercise duration and intensity in a weight-loss program. *Clin J Sport Med*, 15(2):113–5, 2005.

[148] J. M. Jakicic, B. H. Marcus, K. I. Gallagher, M. Napolitano, and W. Lang. Effect of exercise duration and intensity on weight loss in overweight, sedentary women: a randomized trial. *JAMA*, 290(10):1323–30, 2003.

[149] A. M. Knab, R. A. Shanely, K. D. Corbin, F. Jin, W. Sha, and D. C. Nieman. A 45-minute vigorous exercise bout increases metabolic rate for 14 hours. *Medicine & Science in Sports & Exercise*, 43(9):1643–8, 2011.

[150] M. Yoshioka, E. Doucet, S. St-Pierre, N. Almeras, D. Richard, et al. Impact of high-intensity exercise on energy expenditure, lipid oxidation and body fatness. *Int J Obes Relat Metab Disord*, 25(3):332–9, 2001.

[151] G. B. Donato, S. C. Fuchs, K. Oppermann, C. Bastos, and P. M. Spritzer. Association between menopause status and central adiposity measured at different cutoffs of waist circumference and waist-to-hip ratio. *Menopause*, 13(2):280–5, 2006.

[152] M. J. Toth, A. Tchernof, C. K. Sites, and E. T. Poehlman. Effect of menopausal status on body composition and abdominal fat distribution. *Int J Obes Relat Metab Disord*, 24(2):226–31, 2000.

[153] A. R. Genazzani and M. Gambacciani. Effect of climacteric transition and hormone replacement therapy on body weight and body fat distribution. *Gynecol Endocrinol*, 22(3):145–50, 2006.

[154] J. Haarbo, U. Marslew, A. Gotfredsen, and C. Christiansen. Postmenopausal hormone replacement therapy prevents central distribution of body fat after menopause. *Metabolism*, 40(12):1323–6, 1991.

[155] S. R. Davis, C. Castelo-Branco, P. Chedraui, M. A. Lumsden, R. E. Nappi, et al. Understanding weight gain at menopause. *Climacteric*, 15(5):419–29, 2012.

[156] N. Allen, C. Sudlow, P. Downey, T. Peakman, J. Danesh, et al. Uk Biobank: Current status and what it means for epidemiology. *Health Policy and Technology*, 1(3):123–126, 2012.

[157] C. M. Koolhaas, F. J. van Rooij, M. Cepeda, H. Tiemeier, O. H. Franco, and J. D. Schoufour. Physical activity derived from questionnaires and wrist-worn accelerometers: comparability and the role of demographic, lifestyle, and health factors among a population-based sample of older adults. *Clin Epidemiol*, 10:1–16, 2018.

[158] S. Skender, J. Ose, J. Chang-Claude, M. Paskow, B. Bruhmann, et al. Accelerometry and physical activity questionnaires - a systematic review. *BMC Public Health*, 16:515, 2016.

[159] J. F. Sallis and B. E. Saelens. Assessment of physical activity by self-report: status, limitations, and future directions. *Res Q Exerc Sport*, 71, 2000.

[160] K. R. Westerterp. Physical activity as determinant of daily energy expenditure. *Physiol Behav*, 93(4-5):1039–43, 2008.

[161] M. S. Buchowski, K. M. Townsend, K. Y. Chen, S. A. Acra, and M. Sun. Energy expenditure determined by self-reported physical activity is related to body fatness. *Obes Res*, 7, 1999.

[162] A. A. Atienza, R. P. Moser, F. Perna, K. Dodd, R. Ballard-Barbash, et al.

Self-reported and objectively measured activity related to biomarkers using NHANES. *Med Sci Sports Exerc*, 43(5):815–21, 2011.

[163] Y. C. Foong, D. Aitken, T. Winzenberg, P. Otahal, D. Scott, and G. Jones. The association between physical activity and reduced body fat lessens with age - results from a cross-sectional study in community-dwelling older adults. *Exp Gerontol*, 55:107–12, 2014.

[164] T. J. Harris, C. G. Owen, C. R. Victor, R. Adams, et al. A comparison of questionnaire, accelerometer, and pedometer: measures in older people. *Med Sci Sports Exerc*, 41(7):1392–402, 2009.

[165] S. Sabia, P. Cogranne, V. T. van Hees, J. A. Bell, A. Elbaz, et al. Physical activity and adiposity markers at older ages: accelerometer vs questionnaire data. *J Am Med Dir Assoc*, 16(5):438.e7–13, 2015.

[166] M. Wanner, A. Richard, B. Martin, D. Faeh, and S. Rohrmann. Associations between self-reported and objectively measured physical activity, sedentary behavior and overweight/obesity in NHANES 2003-2006. *Int J Obes (Lond)*, 41(1):186–193, 2017.

[167] Y. Kim, T. White, K. Wijndaele, S. J. Sharp, N. J. Wareham, and S. Brage. Adiposity and grip strength as long-term predictors of objectively measured physical activity in 93 015 adults: the UK Biobank study. *Int J Obes (Lond)*, 41(9):1361–1368, 2017.

[168] M. Willetts, S. Hollowell, L. Aslett, C. Holmes, and A. Doherty. Statistical machine learning of sleep and physical activity phenotypes from sensor data in 96,220 uk biobank participants. *Sci Rep*, 8(1):7961, 2018.

[169] L Qi and Y Cho. Gene environment interaction and obesity. *Nutrition reviews*, 66:684–694, 2008.

[170] A E Locke, B Kahali, S I Berndt, A E Justice, T H Pers, et al. Genetic studies of body mass index yield new insights for obesity biology. *Nature*, 518:197–206, 2015.

[171] E K Speliotes, C J Willer, S I Berndt, K L Monda, G Thorleifsson, et al. Association analyses of 249,796 individuals reveal 18 new loci associated with body mass index. *Nat Genet*, 42:937–948, 2010.

[172] T. Rankinen and C. Bouchard. Gene-physical activity interactions: overview of human studies. *Obesity (Silver Spring)*, 16 Suppl 3:S47–50, 2008.

[173] K. S. Vimaleswaran, S. Li, J. H. Zhao, J. Luan, S. Bingham, et al. Physical activity attenuates the body mass index – increasing influence of genetic variation in the FTO gene 1 – 3. *Journal of the American Dietetic Association*, pages 425–428, 2009.

[174] T. O. Kilpelainen, L. Qi, S. Brage, S. J. Sharp, E. Sonestedt, et al. Physical activity attenuates the influence of FTO variants on obesity risk: A meta-analysis of 218,166 adults and 19,268 children. *PLoS Medicine*, 8, 2011.

[175] A. I. Young, F. Wauthier, and P. Donnelly. Multiple novel gene-by-environment interactions modify the effect of FTO variants on body mass index. *Nature Communications*, 7, 2016.

[176] C. F. M. Celis-Morales, C.and Marsaux, K. M. Livingstone, S. Navas-Carretero, R. San-Cristobal, et al. Physical activity attenuates the effect

of the FTO genotype on obesity traits in European adults: The Food4Me study. *Obesity*, 2016.

[177] T. Ahmad, L. E E I-Min, G. Paré, D. I. Chasman, L. Rose, et al. Lifestyle interaction with fat mass and obesity-associated (FTO) genotype and risk of obesity in apparently healthy U.S. women. *Diabetes Care*, 2011.

[178] H. Reddon, H. C. Gerstein, J. C. Engert, V. Mohan, J. Bosch, et al. Physical activity and genetic predisposition to obesity in a multiethnic longitudinal study the confidence in statistical associations. *Nature Publishing Group*, 2015.

[179] A. S. Richardson, K. E. North, M. Graff, K. M. Young, K. L. Mohlke, et al. Moderate to vigorous physical activity interactions with genetic variants and body mass index in a large US ethnically diverse cohort. *Pediatric Obesity*, 2014.

[180] C. H. Andreasen, K. L. Stender-Petersen, M. S. Mogensen, S. S. Torekov, L. Wegner, et al. Low physical activity accentuates the effect of the FTO rs9939609 polymorphism on body fat accumulation. *Diabetes*, 57(1):95–101, 2008.

[181] S. C. Moore, M. J. Gunter, C. R. Daniel, K. S. Reddy, P. S. George, et al. Common genetic variants and central adiposity among Asian-Indians. *Obesity (Silver Spring)*, 20(9):1902–8, 2012.

[182] E. Rampersaud, B. D. Mitchell, T. I. Pollin, M. Fu, H. Shen, et al. Physical activity and the association of common FTO gene variants with body mass index and obesity. *Arch Intern Med*, 168(16):1791–7, 2008.

[183] D. Corella, C. Ortega-Azorin, J. V. Sorli, M. I. Covas, P. Carrasco, et al. Statistical and biological gene-lifestyle interactions of MC4R and FTO with diet and physical activity on obesity: new effects on alcohol consumption. *PLoS One*, 7(12):e52344, 2012.

[184] S. Cauchi, F. Stutzmann, C. Cavalcanti-Proenca, E. Durand, A. Pouta, et al. Combined effects of MC4R and FTO common genetic variants on obesity in European general populations. *J Mol Med (Berl)*, 87(5):537–46, 2009.

[185] J. Zhu, R. J F Loos, L. Lu, G. Zong, W. Gan, et al. Associations of genetic risk score with obesity and related traits and the modifying effect of physical activity in a Chinese Han population. *PLoS ONE*, 2014.

[186] S. Li, J. H. Zhao, J. Luan, U. Ekelund, R. N. Luben, et al. Physical activity attenuates the genetic predisposition to obesity in 20,000 men and women from EPIC-Norfolk prospective population study. *PLoS Medicine*, 7:1–9, 2010.

[187] W Johnson, Kk Ong, Ce Elks, N. Wareham, A Wong, et al. Modification of genetic influences on adiposity between 36 and 63 years of age by physical activity and smoking in the 1946 British birth cohort study. *Nutrition amp; Diabetes*, 4, 2014.

[188] S. Ahmad, G. Rukh, T. V. Varga, A. Ali, A. Kurbasic, et al. Gene x physical activity interactions in obesity: Combined analysis of 111,421 individuals of European ancestry. *PLoS Genetics*, 9:1–9, 2013.

[189] Q. Qi, Y. Li, A. K. Chomistek, J. H. Kang, G. C. Curhan, et al. Television watching, leisure time physical activity, and the genetic predisposition in

relation to body mass index in women and men. *Circulation*, 126:1821–1827, 2012.

[190] M. Graff, K. E. North, A. S. Richardson, K. M. Young, K. L. Mohlke, et al. Screen time behaviours may interact with obesity genes, independent of physical activity, to influence adolescent BMI in an ethnically diverse cohort. *Pediatric Obesity*, 2013.

[191] Y Klimentidis, A Arora, A Chougule, J Zhou, and D Raichlen. FTO association and interaction with time spent sitting. *International Journal of Obesity*, 40:411–416, 2016.

[192] M. Graff, R. A. Scott, A. E. Justice, K. L. Young, M. F. Feitosa, et al. Genome-wide physical activity interactions in adiposity - a meta-analysis of 200,452 adults. *PLoS Genet*, 13(4):e1006528, 2017.

[193] A. Jonsson, F. Renström, V. Lyssenko, E. C. Brito, B. Isomaa, et al. Assessing the effect of interaction between an FTO variant (rs9939609) and physical activity on obesity in 15,925 Swedish and 2,511 Finnish adults. *Diabetologia*, 52:1334–1338, 2009.

[194] J. Y. Kim, J. T. DeMenna, S. Puppala, G. Chittoor, J. Schneider, et al. Physical activity and FTO genotype by physical activity interactive influences on obesity. *BMC Genet*, 17:47, 2016.

[195] T. Huang and F. B. Hu. Gene-environment interactions and obesity: recent developments and future directions. *BMC Medical Genomics*, 8 Suppl 1:S2, 2015.

[196] J. Tyrrell, A. R. Wood, R. M. Ames, H. Yaghootkar, R. N. Beaumont, et al. Gene-obesogenic environment interactions in the UK Biobank study. *Int J Epidemiol*, 46(2):559–575, 2017.

[197] M. Rask-Andersen, T. Karlsson, W. E. Ek, and A. Johansson. Gene-environment interaction study for BMI reveals interactions between genetic factors and physical activity, alcohol consumption and socioeconomic status. *PLoS Genet*, 13(9):e1006977, 2017.

[198] S. Ahmad, W. Zhao, F. Renstrom, A. Rasheed, M. Samuel, et al. Physical activity, smoking, and genetic predisposition to obesity in people from Pakistan: the PROMIS study. *BMC Med Genet*, 16:114, 2015.

[199] H. Reddon, J. L. Gueant, and D. Meyre. The importance of gene-environment interactions in human obesity. *Clin Sci (Lond)*, 130(18):1571–97, 2016.

[200] M. O. Goodarzi. Genetics of obesity: what genetic association studies have taught us about the biology of obesity and its complications. *Lancet Diabetes Endocrinol*, 6(3):223–236, 2018.

[201] M. C. Pike, D. V. Spicer, L. Dahmoush, and M. F. Press. Estrogens, progestogens, normal breast cell proliferation, and breast cancer risk. *Epidemiol Rev*, 15(1):17–35, 1993.

[202] S. Connor Gorber, M. Tremblay, D. Moher, and B. Gorber. A comparison of direct vs. self-report measures for assessing height, weight and body mass index: a systematic review. *Obes Rev*, 8(4):307–26, 2007.

[203] P. H. Lahmann, K. Hoffmann, N. Allen, C. H. van Gils, K. T. Khaw, et al. Body size and breast cancer risk: findings from the European Prospective Investigation into Cancer and Nutrition (EPIC). *Int J Cancer*, 111(5):762–71, 2004.

[204] G. Tagliabue, A. Maghini, S. Fabiano, A. Tittarelli, E. Frassoldi, E. Costa, S. Nobile, T. Codazzi, P. Crosignani, R. Tessandori, and P. Contiero. Consistency and accuracy of diagnostic cancer codes generated by automated registration: comparison with manual registration. *Popul Health Metr*, 4:10, 2006.

[205] S. MacMahon, R. Peto, R. Collins, J. Godwin, J. Cutler, 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*, 335(8692):765–774, 1990.

[206] S. Parish, R. Peto, A. Palmer, R. Clarke, S. Lewington, et al. The joint effects of apolipoprotein B, apolipoprotein A1, LDL cholesterol, and HDL cholesterol on risk: 3510 cases of acute myocardial infarction and 9805 controls. *Eur Heart J*, 30(17):2137–46, 2009.

[207] E. E. Krebs, B. C. Taylor, J. A. Cauley, K. L. Stone, P. J. Bowman, and K. E. Ensrud. Measures of adiposity and risk of breast cancer in older postmenopausal women. *J Am Geriatr Soc*, 54(1):63–9, 2006.

[208] P. H. Lahmann, L. Lissner, B. Gullberg, H. Olsson, and G. Berglund. A prospective study of adiposity and postmenopausal breast cancer risk: the Malmo Diet and Cancer Study. *Int J Cancer*, 103(2):246–52, 2003.

[209] R. J. Macinnis, D. R. English, D. M. Gertig, J. L. Hopper, and G. G. Giles. Body size and composition and risk of postmenopausal breast cancer. *Cancer Epidemiol Biomarkers Prev*, 13(12):2117–25, 2004.

[210] T. E. Rohan, M. Heo, L. Choi, M. Datta, J. L. Freudenheim, et al. Body fat and breast cancer risk in postmenopausal women: a longitudinal study. *J Cancer Epidemiol*, 2013:754815, 2013.

[211] D. W. Losordo and J. M. Isner. Estrogen and angiogenesis: A review. *Arterioscler Thromb Vasc Biol*, 21(1):6–12, 2001.

[212] S. Rinaldi, T. J. Key, P. H. Peeters, P. H. Lahmann, A. Lukanova, et al. Anthropometric measures, endogenous sex steroids and breast cancer risk in postmenopausal women: a study within the EPIC cohort. *Int J Cancer*, 118 (11):2832–9, 2006.

[213] Z. Huang, S. E. Hankinson, G. A. Colditz, M. J. Stampfer, D. J. Hunter, et al. Dual effects of weight and weight gain on breast cancer risk. *JAMA*, 278(17):1407–11, 1997.

[214] M. Harvie, L. Hooper, and A. H. Howell. Central obesity and breast cancer risk: a systematic review. *Obes Rev*, 4(3):157–73, 2003.

[215] M. M. Gaudet, B. D. Carter, A. V. Patel, L. R. Teras, E. J. Jacobs, and S. M. Gapstur. Waist circumference, body mass index, and postmenopausal breast cancer incidence in the Cancer Prevention Study-ii Nutrition Cohort. *Cancer Causes Control*, 25(6):737–45, 2014.

[216] T. A. Sellers, J. Davis, J. R. Cerhan, R. A. Vierkant, J. E. Olson, et al. Interaction of waist/hip ratio and family history on the risk of hormone receptor-defined breast cancer in a prospective study of postmenopausal women. *Am J Epidemiol*, 155(3):225–33, 2002.

[217] M. Song, W. C. Willett, F. B. Hu, D. Spiegelman, A. Must, et al. Trajectory of body shape across the lifespan and cancer risk. *Int J Cancer*, 138(10): 2383–95, 2016.

[218] D. Brodie, V. Moscrip, and R. Hutcheon. Body composition measurement: a review of hydrodensitometry, anthropometry, and impedance methods. *Nutrition*, 14(3):296–310, 1998.

[219] H. C. Lukaski, P. E. Johnson, W. W. Bolonchuk, and G. I. Lykken. Assessment of fat-free mass using bioelectrical impedance measurements of the human body. *Am J Clin Nutr*, 41(4):810–7, 1985.

[220] L. B. Houtkooper, T. G. Lohman, S. B. Going, and W. H. Howell. Why bioelectrical impedance analysis should be used for estimating adiposity. *Am J Clin Nutr*, 64(3 Suppl):436s–448s, 1996.

[221] NIH consensus statement. bioelectrical impedance analysis in body composition measurement. National Institutes of Health Technology assessment conference statement. December 12-14, 1994. *Nutrition*, 12(11-12):749–62, 1996.

[222] J. C. Keen and N. E. Davidson. The biology of breast carcinoma. *Cancer*, 97(3 Suppl):825–33, 2003.

[223] C. M. Friedenreich and A. E. Cust. Physical activity and breast cancer risk: Impact of timing, type and dose of activity and population subgroup effects. *British Journal of Sports Medicine*, 42(8):636–647, 2008.

[224] S. C. Moore, I. M. Lee, E. Weiderpass, P. T. Campbell, J. N. Sampson, et al. Association of leisure-time physical activity with risk of 26 types of cancer in 1.44 million adults. *JAMA Intern Med*, 176(6):816–25, 2016.

[225] S. E. Neil-Sztramko, T. Boyle, E. Milosevic, S. F. Nugent, C. C. Gotay, and K. L. Campbell. Does obesity modify the relationship between physical activity and breast cancer risk? *Breast Cancer Res Treat*, 2017.

[226] M. Plummer. Improved estimates of floating absolute risk. *Stat Med*, 23(1): 93–104, 2004.

[227] A. Bardia, L. C. Hartmann, C. M. Vachon, R. A. Vierkant, A. H. Wang, et al. Recreational physical activity and risk of postmenopausal breast cancer based on hormone receptor status. *Archives of Internal Medicine*, 166(22): 2478–2483, 2006.

[228] S. C. Chang, R. G. Ziegler, B. Dunn, R. Stolzenberg-Solomon, Jr. Lacey, J. V., et al. Association of energy intake and energy balance with postmenopausal breast cancer in the prostate, lung, colorectal, and ovarian cancer screening trial. *Cancer Epidemiol Biomarkers Prev*, 15(2):334–41, 2006.

[229] M. F. Leitzmann, S. C. Moore, T. M. Peters, J. V. Lacey, A. Schatzkin, et al. Prospective study of physical activity and risk of postmenopausal breast cancer. *Breast Cancer Research*, 10(5), 2008.

[230] K. Bhaskaran, I. Douglas, H. Forbes, I. dos Santos-Silva, D. A. Leon, and L. Smeeth. Body-mass index and risk of 22 specific cancers: a population-based cohort study of 5.24 million UK adults. *Lancet*, 384(9945):755–65, 2014.

[231] K. H. Pietilainen, M. Korkeila, L. H. Bogl, K. R. Westerterp, H. Yki-Jarvinen, et al. Inaccuracies in food and physical activity diaries of obese subjects: complementary evidence from doubly labeled water and co-twin assessments. *Int J Obes (Lond)*, 34(3):437–45, 2010.

[232] R. Rzewnicki, Y. Vanden Auweele, and I. De Bourdeaudhuij. Addressing overreporting on the International Physical Activity Questionnaire (IPAQ) telephone survey with a population sample. *Public Health Nutr*, 6(3):299–305, 2003.

[233] S. Sabia, V. T. van Hees, M. J. Shipley, M. I. Trenell, G. Hagger-Johnson, et al. Association between questionnaire- and accelerometer-assessed physical activity: the role of sociodemographic factors. *Am J Epidemiol*, 179(6): 781–90, 2014.

[234] E. A. Spencer, P. N. Appleby, G. K. Davey, and T. J. Key. Validity of self-reported height and weight in 4808 EPIC-Oxford participants. *Public Health Nutr*, 5(4):561–5, 2002.

[235] L. Richiardi, R. Bellocco, and D. Zugna. Mediation analysis in epidemiology: methods, interpretation and bias. *Int J Epidemiol*, 42(5):1511–9, 2013.

[236] P. A. van den Brandt, D. Spiegelman, S. S. Yaun, H. O. Adami, L. Beeson,

et al. Pooled analysis of prospective cohort studies on height, weight, and breast cancer risk. *Am J Epidemiol*, 152(6):514–27, 2000.

[237] A. Amadou, P. Ferrari, R. Muwonge, A. Moskal, C. Biessy, et al. Overweight, obesity and risk of premenopausal breast cancer according to ethnicity: a systematic review and dose-response meta-analysis. *Obesity Reviews*, 14(8): 665–678, 2013.

[238] K. A. Bertrand, T. N. Bethea, L. L. Adams-Campbell, L. Rosenberg, and J. R. Palmer. Differential patterns of risk factors for early-onset breast cancer by ER status in African American women. *Cancer Epidemiol Biomarkers Prev*, 26(2):270–277, 2017.

[239] H. Du, D. Bennett, L. Li, G. Whitlock, Y. Guo, et al. Physical activity and sedentary leisure time and their associations with BMI, waist circumference, and percentage body fat in 0.5 million adults: the China Kadoorie Biobank study. *Am J Clin Nutr*, 97(3):487–96, 2013.

[240] B. Lauby-Secretan, C. Scoccianti, D. Loomis, Y. Grosse, F. Bianchini, and K. Straif. Body fatness and cancer–viewpoint of the IARC Working Group. *N Engl J Med*, 375(8):794–8, 2016.

[241] A. J. White, H. B. Nichols, P. T. Bradshaw, and D. P. Sandler. Overall and central adiposity and breast cancer risk in the Sister Study. *Cancer*, 121(20): 3700–8, 2015.

[242] The Endogenous Hormones and Breast Cancer Collaborative Group. Endogenous sex hormones and breast cancer in postmenopausal women: reanalysis of nine prospective studies. *J Natl Cancer Inst*, 94(8):606–16, 2002.

[243] T. J. Key. Endogenous oestrogens and breast cancer risk in premenopausal and postmenopausal women. *Steroids*, 76(8):812–5, 2011.

[244] F. Islami, Y. Liu, A. Jemal, J. Zhou, E. Weiderpass, et al. Breastfeeding and breast cancer risk by receptor status–a systematic review and meta-analysis. *Ann Oncol*, 26(12):2398–407, 2015.

[245] R. G. Ziegler, R. N. Hoover, M. C. Pike, A. Hildesheim, A. M. Nomura, et al. Migration patterns and breast cancer risk in Asian-American women. *J Natl Cancer Inst*, 85(22):1819–27, 1993.

[246] J. J. Varo, M. A. Martinez-Gonzalez, J. De Irala-Estevez, J. Kearney, M. Gibney, and J. A. Martinez. Distribution and determinants of sedentary lifestyles in the European Union. *Int J Epidemiol*, 32(1):138–46, 2003.

[247] S. Rinaldi, R. Kaaks, C. M. Friedenreich, T. J. Key, R. Travis, et al. Physical activity, sex steroid, and growth factor concentrations in pre- and postmenopausal women: a cross-sectional study within the EPIC cohort. *Cancer Causes Control*, 25(1):111–24, 2014.

[248] J. West, O. Dahlqvist Leinhard, T. Romu, R. Collins, S. Garratt, et al. Feasibility of mr-based body composition analysis in large scale population studies. *PLoS One*, 11(9):e0163332, 2016.

[249] M. Borga, J. West, J. D. Bell, N. C. Harvey, T. Romu, et al. Advanced body composition assessment: from body mass index to body composition profiling. *J Investig Med*, 2018.