Explaining the mechanisms underlying the effect of adiposity on the risk of colorectal, postmenopausal breast, and endometrial cancer using data from four prospective cohort studies

Seyedeh Ghazaleh Dashti

ORCID: orcid.org/0000-0002-1399-7220

Submitted in total fulfilment of the requirements of the degree of

Doctor of Philosophy

Centre for Epidemiology and Biostatistics

Melbourne School of Population and Global Health

THE UNIVERSITY OF MELBOURNE

July 2020

Abstract

Background

Excess adiposity in adulthood is an established risk factor for several types of cancer, including colorectal, estrogen receptor (ER)-positive postmenopausal breast, and endometrial cancers. Dysregulated adipocytokine production and chronic low-grade inflammation, insulin resistance and hyperinsulinemia, and alteration in levels of sex-steroid hormones are three of the mechanisms hypothesised to explain the effect of adiposity on cancer risk. Current evidence suggests that these mechanisms are interconnected, and their relative importance is likely different for different cancers.

Mediation analysis is a statistical method for quantifying how much of the effect of an exposure (adiposity) on an outcome (cancer incidence) is explained by a mediator (an intermediate mechanism). Recent methodological advances in this area allow assessment of mediation in the presence of multiple dependent mediators and are of great use in unravelling mechanisms linking adiposity to cancer risk. This thesis took advantage of these methods and data from four existing prospective studies to provide a comprehensive investigation of the role of the three putative pathways in explaining the effect of adiposity on the risk of colorectal, ER-positive postmenopausal breast, and endometrial cancers.

Methods

The UK Biobank cohort was used to investigate the mediating roles of measured biomarkers on inflammatory status (C-reactive protein (CRP)), insulin signalling (glycated haemoglobin), and sex steroid hormone (sex hormone binding globulin (SHBG) and testosterone) pathways in explaining the effect of adiposity on colorectal cancer risk in men and postmenopausal women. Sequential mediation analysis was used

to estimate the natural indirect (NIE) and direct (NDE) effects. A case-cohort study within the Melbourne Collaborative Cohort Study (MCCS) was used to investigate the mediating roles of fasting insulin and calculated free estradiol in explaining the effect of adiposity on ER-positive postmenopausal breast cancer risk. A regression-standardisation approach was used to estimate the interventional indirect (IIE) and direct (IDE) effects. A case-control study nested within the European Prospective Investigation into Cancer and Nutrition (EPIC) was used to investigate the mediating roles of biomarkers on inflammatory status (captured by a panel of pro- and anti-inflammatory adipocytokines and cytokines), C-peptide, and estrone and calculated free estradiol in explaining the effect of adiposity on endometrial cancer risk in postmenopausal women. The NIEs and NDE were estimated using sequential mediation analysis. Finally, a case-cohort study within the Women's Health Initiative Observational Study (WHI-OS) was used to investigate the mediating roles of biomarkers on inflammatory status (a panel of pro- and anti-inflammatory adipocytokines and cytokines), fasting insulin, and total estradiol in explaining the effect of adiposity on the risk of colorectal, ER-positive breast, and endometrial cancers in postmenopausal women. The IIEs and IDE were estimated using a regression-standardisation approach.

Results

The analysis of the UK Biobank data provided evidence for a small mediating role for CRP as well as for SHBG and testosterone (combined and beyond the potential influences of CRP and HbA1c on SHBG and testosterone levels), but no evidence of mediation through HbA1c in explaining the effect of adiposity on colorectal cancer risk in men. In postmenopausal women, there was evidence for a small mediating effect through CRP, but no evidence for mediation through HbA1c or SHBG and testosterone

combined. For both men and women, a large part of the effect of adiposity on colorectal cancer risk remained unexplained by the biomarkers included in the analysis.

In the MCCS case-cohort study of adiposity and ER-positive postmenopausal breast cancer, moderate to strong mediation was observed through free estradiol. The point estimate was also suggestive of weak mediation through fasting insulin, but there was high uncertainty around the estimated indirect effect.

In the case-control study nested within the EPIC there was evidence for moderate mediation through inflammatory status. There was also a suggestion for weak mediation through C-peptide (excluding the potential influence of inflammatory status on C-peptide levels) and weak to moderate mediation through free estradiol and estrone (excluding the influences of inflammatory status and C-peptide on sex-hormone levels), but there was high uncertainty around the estimated indirect effects.

In the WHI-OS case-cohort study of postmenopausal women, for colorectal cancer there was high uncertainty around the estimated mediating effects for all pathways, precluding the possibility of making a definitive judgement about presence and strength of mediation through the assessed biomarkers. For ER-positive breast cancer, a strong mediating effect was observed through fasting insulin and a weak to moderate mediation for estradiol, but there was high uncertainty around the estimated indirect effect through inflammatory status. For endometrial cancer, a moderate mediating effect through estradiol was observed, but there was high uncertainly around the estimated indirect effects for inflammatory status and fasting insulin.

Conclusion

This thesis offers an investigation of the mediating roles of inflammatory status, insulin resistance and hyperinsulinemia, and sex-steroid hormone pathways in explaining the effect of adiposity on the risk of colorectal cancer in men and postmenopausal women, and ER-positive breast and endometrial cancer in postmenopausal women. Findings from this research need to be replicated by studies ideally with larger sample sizes and repeated measures for adiposity and biomarkers, and as has been performed in this thesis, proper application of causal mediation analysis approaches to take the dependence between biomarkers into account.

Declaration

This is to certify that:

1. the thesis comprises only my original work towards the Doctor of Philosophy, except where indicated in the Preface;

2. due acknowledgement has been made in the text to all other material used;

3. the thesis is less than 100,000 words in length, exclusive of tables, maps, bibliographies, and appendices.

Acknowledgements

On Friday 21st June 2019, around 7:30 pm, I wrote in my diary "*I am looking forward to writing the acknowledgements of my thesis*". I wrote this in Farsi, the language I first learnt to communicate in and still use for my diaries. On the day, I remember I was on a Melbourne tram (no 75), half-listening to a podcast, and feeling somewhat irritated at a non-converging multiple imputation model. Yes, looking forward to writing my thesis acknowledgements was a random thought… Back then, I still had three analyses to complete, three manuscripts to write, and I had not even started to think about the structure of the thesis. Did I mention the non-converging multiple imputation model?

Like everyone else, I went through my PhD journey without being able to stop life from happening. On the good days, I felt fortunate for being surrounded by the most caring and wisest people. On the bad days, I was even more aware of how blessed I was for having the support of those important individuals. So, here I am, nine months after that tram ride, relishing the opportunity to thank the key people without whom the completion of this PhD would have been impossible.

Above all, I would like to express my sincerest gratitude to my supervisors: Professor Dallas English, Professor Julie Simpson, Dr Amalia (Emily) Karahalios, and Dr Vivian Viallon. Thank you all for being the dream supervisory team and the perfect role models, for your time, ongoing support, all the opportunities you have given me, and for your immense kindness.

Dallas introduced me to and instilled in me the love for epidemiology during my master's degree. I am so grateful to him for agreeing to be my primary PhD supervisor. He is the most tremendous mentor and supervisor: knowledgeable, reliable, frank,

open-minded, and humble. I am thankful for his unwavering support and kindness throughout my candidature. Julie's attributes include her formidable intelligence and superb teaching. The graphs and formulae she quickly sketched on the notepad, a tool that was always sitting on her desk, helped get me through the challenging statistical aspects of this thesis. She inspires me every day, and I am grateful to her for her mentorship and for sharing her wisdom and experiences. Emily has also been a wonderful supervisor, unquestionably enthusiastic, always kind, and insightful with eagle eyes. I treasure all our conversations and the advice she has given me about PhD, life, and everything in between. In France, Vivian took me under his wing when I went to Lyon to spend nine months at the International Agency for Research on Cancer (IARC). Neither his academic input and support nor, indeed, his sense of humour diminished upon my return to Melbourne.

I am also obliged to Dr Margarita Moreno Betancur and Dr Marc Gunter for their significant contributions to this thesis, and for giving me their time and sharing their knowledge. I would also like to thank Dr Neil Murphy and Dr Laure Dossus for their generosity and support. Further, I must sincerely acknowledge the contribution of my advisory committee comprising Professor Shyamali Dharmage and Professor Tony Blakely. Separately, Professor Fiona Stanley and Professor Rob Moodie have been wonderful mentors to me.

I was fortunate to spend nine fruitful and memorable months in France working at IARC. I would like to thank Dr Pietro Ferrari and Dr Marc Gunter there for the opportunity and for welcoming me into their group. I would also like to thank all the amazing people at the Nutritional Methodology and Biostatistics Group, who made my stay in Lyon incredibly pleasant.

Thank you also to everyone at the Centre for Epidemiology and Biostatistics at Melbourne School of Population and Global Health. This place has become my second home in my adopted city of Melbourne, and I will always cherish the friendships developed during my time at this centre.

On a personal note, I would like to thank my dearest friends who have all brightened my days, warmed my heart, and kept me sane over the close friendships we've shared. Finally, thank you to the people closest to my heart, Mum, Dad, Sam, and Navid. They have each always been there for me, encouraged me, supported me, and cheered for me, no matter what. I am extremely lucky to have had their loving support over the years. I know I would have not been able to achieve what I have in my life - including completion of this PhD - if not for their contribution.

Preface

I was supported by a Melbourne Research Scholarship (from October 2016 to December 2017 and from October 2019 to April 2020), and a Postgraduate Scholarship from the Australian National Health and Medical Research Council (NHMRC, 1150591; from January 2018 to September 2019). I also undertook part of this research while hosted by the International Agency for Research on Cancer (IARC) and funded partially by a PhD Scholarship by IARC (from May 2018 to February 2019).

This thesis uses data from the UK Biobank (presented in Chapter 5), the Melbourne Collaborative Cohort Study (Chapter 6), the European Prospective Investigation into Cancer and Nutrition (Chapter 7), and the Women's Health Initiative-Observational Study (Chapter 8). Acknowledgement of financial support for each study is provided at the end of the relevant chapters.

Chapter 5 of this thesis includes the author accepted version of the following paper*:* "Dashti SG, Viallon V, Simpson JA, Karahalios A, Moreno-Betancur M, English DR, Gunter MJ, Murphy N. Explaining the link between adiposity and colorectal cancer risk in men and postmenopausal women in the UK Biobank: a sequential causal mediation analysis. Int J Cancer. (1097-0215 (Electronic)). DOI: 10.1002/ijc.32980". This paper is in production for publication at the time of thesis submission, and has been included in the thesis in accordance with *Wiley's Article Sharing Guidelines* available at https://authorservices.wiley.com/asset/Article_Sharing_Guidelines.pdf. For this paper, I planned and performed the data analysis, interpreted the results, wrote the initial manuscript draft, and revised it following feedback from the co-authors and journal reviewers.

Chapter 6 of this thesis includes the author accepted version of the following publication "Dashti SG, Simpson JA, Karahalios A, Viallon V, Moreno-Betancur M, Gurrin LC, MacInnis RJ, Lynch BM, Baglietto L, Morris HA, Gunter MJ, Ferrari P, Milne RL, Giles GG, English DR. Adiposity and estrogen receptor-positive, postmenopausal breast cancer risk: Quantification of the mediating effects of fasting insulin and free estradiol. Int J Cancer. 2020;146(6):1541-52. DOI: 10.1002/ijc.32504". As above, this paper has been included in the thesis in accordance with *Wiley's Article Sharing Guidelines.* For this paper, I helped with the data cleaning, planned and performed the data analysis, interpreted the results, wrote the initial manuscript draft, and revised it following feedback from the co-authors and journal reviewers.

Chapter 7 of this thesis includes the manuscript submitted to Cancer Epidemiology, Biomarkers & Prevention. Maintaining the confidentiality of the findings in this paper, is kindly requested. For this paper, I planned and performed the data analysis, interpreted the results, wrote the initial manuscript draft, and revised it following feedback from the co-authors.

Chapter 8 of this thesis includes the draft of an in-progress manuscript. The work presented here was conducted after the proposal for the analysis was approved by the WHI Publications and Presentations (P&P) committee (proposal number 3809). The proposal authors are S. Ghazaleh Dashti, Marc J. Gunter, Gloria Y.F. Ho, Howard D. Strickler, and Sylvia Wassertheil-Smoller. Professor Sylvia Wassertheil-Smoller is also the sponsoring WHI Principal Investigator for the proposal. Julie A Simpson, Vivian Viallon, Amalia Karahalios, Margarita Moreno-Betancur, Dallas R English, and Marc J Gunter have read and contributed to the version of the manuscript included in the thesis. However, this is not the final author list and it will be expanded prior to submission of

the manuscript to a nominated journal. Maintaining the confidentiality of the findings in this paper is kindly requested. For this paper, I planned and performed the data analysis, interpreted the results, wrote the initial manuscript draft, and revised it following feedback from the co-authors listed above.

Parts of Chapters 1 to 3 of this thesis have been adapted from submitted grant proposals to the NHMRC (APP1138594, APP1163513), World Cancer Research Fund (WCRF) International Regular Grant Programme (2018/1730), and my NHMRC postgraduate scholarship proposal (1150591). On the three project grants, I was listed as a co-investigator and contributed to writing the proposals.

Spelling, referencing, tables, and figure numbers have been modified across all chapters for consistency. Throughout the thesis, the Australian spelling has been used, except for estrogen, estradiol, and estrone, where the American spelling has instead been used.

My advisory committee approved the inclusion of the published and in progress works outlined above.

Contents

Chapter 1: Introduction 1
1.1 Background 1
1.2 Problem statement 4
1.3 Aim and scope 5
1.4 Significance of the thesis 6
Chapter 2: Background 7
2.1 Adiposity 7
2.2 Adiposity and cancer 13
2.2.1 Colorectal cancer 14
2.2.2 Breast cancer 17
2.2.3 Endometrial cancer 19
2-3 Mechanisms linking adiposity to cancer 22
2.4 Summary 62
Chapter 3: Mediation analysis 64
3.1 Setting of mediation analysis 65
3.2 Traditional approaches to mediation analysis 66
3.3 Causal mediation analysis 68
3.4 No confounding assumptions for mediation analysis 73
3.5 Estimation of the effects under causal mediation analysis 74
3.6 Proportion mediated 78
3.7 Mediation analysis with multiple mediators 79
3.7.1 Joint analysis 80
3.7.2 Sequential analysis 81
3.7.3 Interventional analysis 84
3.8 Inclusion of interaction in mediation analysis 90
3.9 Epidemiological studies investigating mediating roles of biomarkers in the effect of adiposity on cancer risk 91
3.10 Summary 102
Chapter 4 Methods 104
4.1 Studies 104
4.2 Statistical analysis 114
Chapter 5: Adiposity and colorectal cancer risk in men and postmenopausal women in the UK Biobank 119
Chapter 6: Adiposity and estrogen receptor-positive breast cancer risk in postmenopausal women in the MCCS 151
Chapter 7: Adiposity and endometrial cancer risk in postmenopausal women in the EPIC 178
Chapter 8: Adiposity and colorectal, estrogen receptor-positive breast, and endometrial cancer risk in postmenopausal women in the WHI-OS 203
Chapter 9: Discussion 235
9.1 Summary of the main findings 235
9.2 Strengths and limitations 242
9.3 Discussion and implications of the findings 246
9.4 Future research 251
9.5 Conclusions 253
References 255
Appendices 266
Appendix A Supplementary material related to publication presented in Chapter 5 267
Appendix B Supplementary material related to publication presented in Chapter 6 285
Appendix C Supplementary material related to publication presented in Chapter 7 288
Appendix D Supplementary material related to publication presented in Chapter 8 296

List of Figures

Figure 1-1 Schematic outline of thesis 6

Figure 2-1 Summary of mechanisms hypothesised to underlie the adiposity-cancer link 24
Figure 2-2 Diagram illustrating the three major metabolic and endocrinological pathways hypothesised to link adiposity to cancer incidence and their postulated interrelations. 25

Figure 3-1 Mediation of the effect of an exposure on an outcome 65
Figure 3-2 Schematic illustration of the counterfactual definition of the total effect of adiposity, as a binary variable (with vs without obesity), on cancer incidence 70
Figure 3-3 Schematic illustration of the counterfactual definition of the total effect of adiposity, as a binary variable (with vs without obesity), on cancer incidence 70
Figure 3-4 Schematic illustration of the counterfactual definition of the natural indirect effect of adiposity, as a binary variable (with vs without obesity), on cancer incidence through estradiol levels 71
Figure 3-5 Schematic illustration of the counterfactual definition of the natural direct effect of adiposity, as a binary variable (with vs without obesity), on cancer incidence not through estradiol levels 72
Figure 3-6 Schematic illustration of the counterfactual definition of the controlled direct effect of adiposity, as a binary variable (with vs without obesity), on cancer incidence when estradiol level is fixed at level $\boldsymbol{m}$ 73
Figure 3-7 Causal diagram illustrating the no-cofounding assumptions required for mediation analysis 74
Figure 3-8 Diagram illustrating the violation of no mediator-outcome confounder affected by the exposure in the presence of multiple mediators affecting each other 79
Figure 3-9 Diagram illustrating estimating indirect and direct effects with $\boldsymbol{M}$ representing the vector of mediators $\boldsymbol{M1}$, $\boldsymbol{M2}$, $\boldsymbol{M3}$ 81
Figure 3-10 Diagrams illustrating the sequential mediation analysis approach 82
Figure 3-11 Schematic illustration of the definition of the interventional indirect effect of adiposity, as a binary variable (with vs without obesity), on cancer incidence through estradiol 86
Figure 3-12 Schematic illustration of the definition of the interventional direct effect of adiposity, as a binary variable (with vs without obesity), on cancer incidence not through estradiol 87
Figure 3-13 Diagrams illustrating the interventional indirect effects estimated using the decomposition approach proposed by Vansteelandt and Daniel (130) 89

Figure 5-1 Flow diagram demonstrating the selection process of men in the UK Biobank for the present study 128
Figure 5-2 Flow diagram demonstrating the selection process of women in the UK Biobank for the present study 129
Figure 5-3 Schematic interpretation of the estimated total, natural direct, and natural indirect effects In this study, C-reactive protein (CRP) was used as a general marker for inflammation, HbA1C as a marker for glycaemic control and a correlate of insulin resistance, and sex hormone binding globulin (SHBG) and testosterone treated jointly as makers of the sex-steroid hormone pathway. Waist circumference, BMI, and waist-hip ratio were used as proxy measures for adiposity in separate sets of analyses. In each provided outcome model needed to estimate the effects, y is the outcome (colorectal cancer incidence), a is the exposure (each adiposity measure comparing obese vs normal), c confounders, m1 CRP, m2 HbA1c, m3a SHBG, and m3b testosterone 134

Figure 6-1 Flowchart of female participants in the Melbourne Collaborative Cohort Study; abbreviation: ER estrogen receptor 159
Figure 6-2 Assumed causal diagram for quantifying the mediating effects of fasting insulin, and free estradiol on adiposity- estrogen receptor-positive postmenopausal breast cancer association; abbreviations: ER estrogen receptor 162

Figure 7-1 Assumed causal structure underlying the effect of adiposity on endometrial cancer 187
Figure 7-2 Interpretation of natural indirect and direct effects*; the solid arrows represent pathways estimated under each effect. 190

Figure 8-1 Assumed causal diagram for quantifying the mediating effects of inflammatory status, fasting insulin, and estradiol on adiposity-cancer (estrogen receptor positive breast, endometrial, and colorectal cancer) association in postmenopausal women 213

List of Tables

Table 2-1 The World Health Organisation suggested thresholds for body mass index, waist circumference, and waist-hip ratio for adults....8
Table 2-2 Pearson's correlation coefficients between measured body mass index, waist circumference, and waist-hip ratio for participants in two large prospective cohort studies also used in this thesis, the UK Biobank and the Melbourne Collaborative Cohort Study....9
Table 2-3 Summary of the WCRF Continuous Update Project dose-response meta-analyses for the association between adiposity and risk of colorectal cancer (34)....16
Table 2-4 Summary of the associations between adiposity measures and colorectal cancer risk estimated in Mendelian randomisation studies....17
Table 2-5 Summary of the WCRF Continuous Update Project dose-response meta-analyses for the association between adiposity and risk of postmenopausal breast cancer (41)....19
Table 2-6 Summary of the associations between adiposity measures and postmenopausal breast cancer risk estimated in Mendelian randomisation studies....19
Table 2-7 Summary of the WCRF Continuous Update Project dose-response meta-analyses for the association between adiposity and risk of endometrial cancer (45, 47)....21
Table 2-8 Summary of the associations between adiposity measures and endometrial cancer risk estimated in Mendelian randomisation studies....21
Table 2-9 Associations for colorectal cancer risk for biomarkers on disturbed adipocytokine production and chronic low-grade inflammation pathway....30
Table 2-10 Associations for breast cancer risk for biomarkers on disturbed adipocytokine production and chronic low-grade inflammation pathway....34
Table 2-11 Associations for endometrial cancer risk for biomarkers on disturbed adipocytokine production and chronic low-grade inflammation pathway....38
Table 2-12 Associations for colorectal cancer risk for biomarkers on insulin resistance and hyperinsulinemia pathway....43
Table 2-13 Associations for breast cancer risk for biomarkers on insulin resistance and hyperinsulinemia pathway....47
Table 2-14 Associations for endometrial cancer risk for biomarkers on insulin resistance and hyperinsulinemia pathway....50
Table 2-15 Associations for colorectal cancer risk for biomarkers on alteration in levels of sex-steroid hormones pathways....54
Table 2-16 Associations for breast cancer risk for biomarkers on alteration in levels of sex-steroid hormones pathways....57
Table 2-17 Associations for endometrial cancer risk for biomarkers on alteration in levels of sex-steroid hormones pathways....59

Table 3-1 Summary of causal mediation analysis approaches for effect decomposition in the presence of multiple interrelated mediators....90
Table 3-2 Total, direct, and indirect effects for the association between adiposity and colorectal, colon, or rectal cancers estimated from cohort, nested case-control or case cohort studies....93
Table 3-3 Total, direct, and indirect effects for the association between adiposity and postmenopausal breast cancer estimated from cohort, nested case-control or case cohort studies....96
Table 3-4 Total, direct, and indirect effects for the association between adiposity and endometrial cancer estimated from cohort, nested case-control or case cohort studies....100

Table 4-1 Overview of the characteristics of the four cohort studies included in this thesis....106
Table 4-2 Designs of the sub-studies within each cohort this thesis made use of, the cancer outcome the studies contributed and their end of follow-up, and the eligibility criteria imposed on each study by the analyses presented in this thesis....109
Table 4-3 Available biomarker measurements in each study that were considered as potential mediators of the effect of adiposity on cancer risk....111
Table 4-4 Overview of the statistical analyses employed for each study....117

Table 5-1 Baseline Characteristics of men and women eligible for the study by categories of waist circumference....137
Table 5-2 Association between adiposity measures and biomarkers; multiple imputation analyses excluding women in the middle category of each adiposity measure....139

Table 5-3 Estimated odds ratios and 95% confidence intervals for the association between adiposity measures (exposure) or biomarkers (mediators) and colorectal cancer (outcome) in men and women; multiple imputation analysis 141
Table 5-4 Estimated natural direct and indirect effects using sequential mediation analysis for the association between adiposity measures and colorectal cancer for men and women, analyses exclude participants categorized as overweight and do not include any interaction terms between adiposity measures and biomarkers (exposure-mediator interactions) or biomarkers (mediator-mediator interactions) - multiple imputation analysis 143

Table 6-1 The interpretations of interventional direct and indirect effects estimated in this study 164
Table 6-2 Baseline characteristics of cases and controls 167
Table 6-3 Estimated relative change in geometric mean of fasting insulin and calculated free estradiol levels associated with measures of adiposity in a random sample of the cohort eligible for this study and available for complete-case analysis (sub-cohort members; n=649; 9 with breast cancer) 168
Table 6-4 Association between measures of adiposity, fasting insulin, and calculated free estradiol and ER-positive postmenopausal breast cancer risk 169
Table 6-5 Results of multiple-mediator analysis with fasting insulin and free estradiol as the mediators of interest, body mass index (≥30 vs <25 kg/m^2) or waist circumference (>88 vs ≤80 cm) as exposure, and non-estrogen receptor negative postmenopausal breast cancer as outcome. 171

Table 7-1 Baseline characteristics of endometrial cancer cases and controls 192
Table 7-2 Exposure-mediator association; estimated ratios in geometric mean of biomarkers associated with body mass index and waist circumference; multiple imputation analyses limited to controls and excluding women categorized as overweight 193
Table 7-3 Estimated exposure-outcome and mediator-outcome associations; multiple imputation analysis 195
Table 7-4 Estimated natural direct and indirect effects using sequential mediation analysis; analyses excluded women categorized as overweight; multiple imputation analysis 197

Table 8-1 Number of women affected and unaffected with endometrial, ER-positive breast, and colorectal cancer in the original case-cohort study, number to be excluded, and the final number in the analytic dataset for this study 211
Table 8-2 Interpretations of the interventional indirect and direct effects 217
Table 8-3 Baseline characteristics of women included in the analytic dataset for mediation analysis 219
Table 8-4 Adiposity-biomarker association; analysis limited to sub-cohort members who met the eligibility criteria and excluding women categorised as overweight 221
Table 8-5 Exposure-outcome and mediator-outcome associations 223
Table 8-6 Interventional indirect and direct effects 226

Table 9-1 Estimated indirect and direct effects for the association between adiposity measures and colorectal cancer risk for men and postmenopausal women 237
Table 9-2 Estimated indirect and direct effects for the association between adiposity measures and estrogen receptor positive breast cancer for postmenopausal women 239
Table 9-3 Estimated indirect and direct effects for the association between adiposity measures and endometrial cancer for postmenopausal women 241

Chapter 1: Introduction

1.1 Background

There is convincing evidence that excess adiposity (overweight and obesity) in adulthood increases the risk of several types of cancer, including colorectal, estrogen receptor (ER)-positive postmenopausal breast, and endometrial cancers (1, 2). Excess adiposity already makes a considerable contribution to the burden of these cancers. In 2012, about 13% of colon and 6% of rectal cancer cases diagnosed in men, 8% of colon and 4% of rectal cancer cases in women, 10% of postmenopausal breast cancer cases, and 34% of endometrial cancer cases were attributable to overweight and obesity (body mass index (BMI) $\geq$25 kg/m^2) (3). Worldwide, the prevalence of obesity has been increasing, and no region has been able to reverse this trend over the past decades (4). This prevalence is predicted to increase further (4), leading to an expected greater contribution of obesity to the burden of cancer in the future.

The mechanisms by which adiposity causes cancer are uncertain, but three major metabolic and endocrinological alterations, known to be driven by adiposity, are likely candidates (1). These include disturbed adipocytokine production and chronic low-grade inflammation, insulin resistance and hyperinsulinemia, and alteration in levels of sex-steroid hormones (1).

Inflammation is a known effect of adiposity (1). In the obese state, pro-inflammatory adipocytokines and cytokines are secreted at higher levels, while the production of adiponectin (an anti-inflammatory adipocytokine) is decreased (1). This inflammatory status can promote carcinogenesis by enabling several hallmarks of cancer, including cell proliferation, apoptosis, and angiogenesis (5). It can also impair insulin signalling,

leading to insulin resistance and hyperinsulinemia (6-9), and enhance aromatase expression, thereby increase estrogen production by the adipose tissue (10). In men with severe obesity, inflammation might contribute to decreased testosterone levels through suppression of the hypothalamic-pituitary-testicular (HPT axis) (11). Finally, sex hormone binding globulin (SHBG) levels might be reduced as a result of inflammation in the obese state, (12) which in turn would increase levels of bioavailable estrogens and testosterone.

Insulin resistance, and as a result hyperinsulinemia, are more common in individuals with obesity (8). Insulin can influence cancer risk via its mitogenic and anti-apoptotic properties (1, 13, 14) and increasing levels of bioavailable insulin-like growth factor (IGF)-1, which also has growth promoting effects (14, 15). Hyperinsulinemia might also lead to increased estrogen and bioavailable estrogen levels through enhancing aromatase activation (10) and downregulating SHBG production (12).

In postmenopausal women with obesity, SHBG levels are reduced, while estrogen and androgen levels are increased (16). Similarly, in men, SHBG levels are higher in the obese state (17), but, contrary to women, total testosterone level is reduced (11). Estrogens promote proliferation of breast and endometrial tissues (1), while they might have antiproliferative effects on colon tissue through their interaction with estrogen receptor-β (ERβ), which is expressed in the colonic epithelium (1). Androgens might also influence cancer risk through their inhibitory effects on colorectal cancer tumour growth (18).

A better understanding of the role of these pathways in mediating the effect of adiposity on cancer risk might help to identify novel risk-reducing interventions for individuals with obesity. Observational epidemiological studies can contribute to this knowledge by

measuring relevant biomarkers reflecting these pathways and statistical mediation analyses of these biomarker data can quantify their contribution to the effect of adiposity on cancer risk (19). A mediation analysis approach most commonly used in studies attempting such an investigation is known as the difference method (19). This approach involves estimating the adiposity-cancer association first without, then with the biomarker (mediator) in the regression model (19). Attenuation of the measure of association after including the mediator is interpreted as evidence of mediation. A limitation of the difference method is that it does not accommodate interactions between the effect of the exposure (adiposity) and mediator on the outcome (cancer incidence) (19), nor does it allow nonlinear exposure-outcome or mediator-outcome associations. Moreover, this approach does not extend to situations where multiple mediators are affecting or interacting with each other (19). As described above, in the case of adiposity and cancer risk, it is highly unlikely that the pathways (mediators) connecting the two operate in isolation. Ignoring dependencies between involved mediators could lead to invalid conclusions about the presence and strength of mediation (19). It is, therefore, imperative to use proper mediation analysis methods that appropriately account for dependencies between pathways.

Recently, mediation analysis has undergone a methodological overhaul based on the counterfactual view of causation (19). Under this definition, for a binary exposure, an individual has two counterfactual outcomes, outcome when exposed, and outcome had the individual been unexposed "with everything else remaining the same", although only one outcome is observable. The resulting framework for causal mediation analysis overcomes the limitations of the difference method and extensions of this framework are now available that allow assessment of mediation in the presence of multiple dependent mediators (19).

Despite these methodological advances, few epidemiological studies have investigated the mediating role of the three putative pathways linking adiposity to cancer risk simultaneously, and even fewer have attempted to disentangle these pathways to assess their potentially different mediating roles. This is the motivation for this thesis. This thesis focuses on providing a more comprehensive investigation of the role of these pathways in explaining the effect of adiposity on the risk of colorectal, ER-positive postmenopausal breast, and endometrial cancers, using data from existing studies and novel causal mediation analysis approaches to analyse the data.

1.2 Problem statement

In 2012, the Director of the US National Cancer Institute listed several provocative questions aimed to engage researchers in addressing major unsolved questions in cancer research. One of these questions concerned mechanisms explaining how adiposity affects cancer risk. His article in Nature stated: "Uncovering the molecular mechanisms by which this happens could help to identify new approaches to cancer prevention or treatment. Such an effort would be a powerful example of how to link risk identification to the molecular steps that drive carcinogenesis" (20). This problem is still unsolved. In a major review of obesity and cancer in 2016, Nimptsch and Pischon (13) stated: "more studies investigating various implicated obesity-associated biomarkers simultaneously in order to quantify their individual mediating role will pave the way for targeted pharmacological or lifestyle interventions aiming at obesity-associated cancer prevention through modification of the most important biomarkers" (p. 210). Identifying intervention targets based on mechanisms is a key component of precision medicine and public health. This approach complements population-wide prevention of weight gain and maintenance of a healthy weight. It also offers hope to people with obesity, most of whom have great difficulty in losing weight and maintaining weight loss (21).

Understanding how adiposity causes cancer and identifying the mechanistic pathways that contribute most to the disease might provide important information for developing and testing strategies to reduce the risk of disease in individuals with obesity while not causing harm (13).

1.3 Aim and scope

The aim of this thesis is to investigate the three pathways that are hypothesised to mediate the relationship between adiposity and colorectal cancer risk in men and postmenopausal women, and adiposity and ER-positive breast and endometrial cancer risk in postmenopausal women. Specifically, this thesis attempts to estimate how much of the effect of adiposity on the risk of these cancers is mediated by:

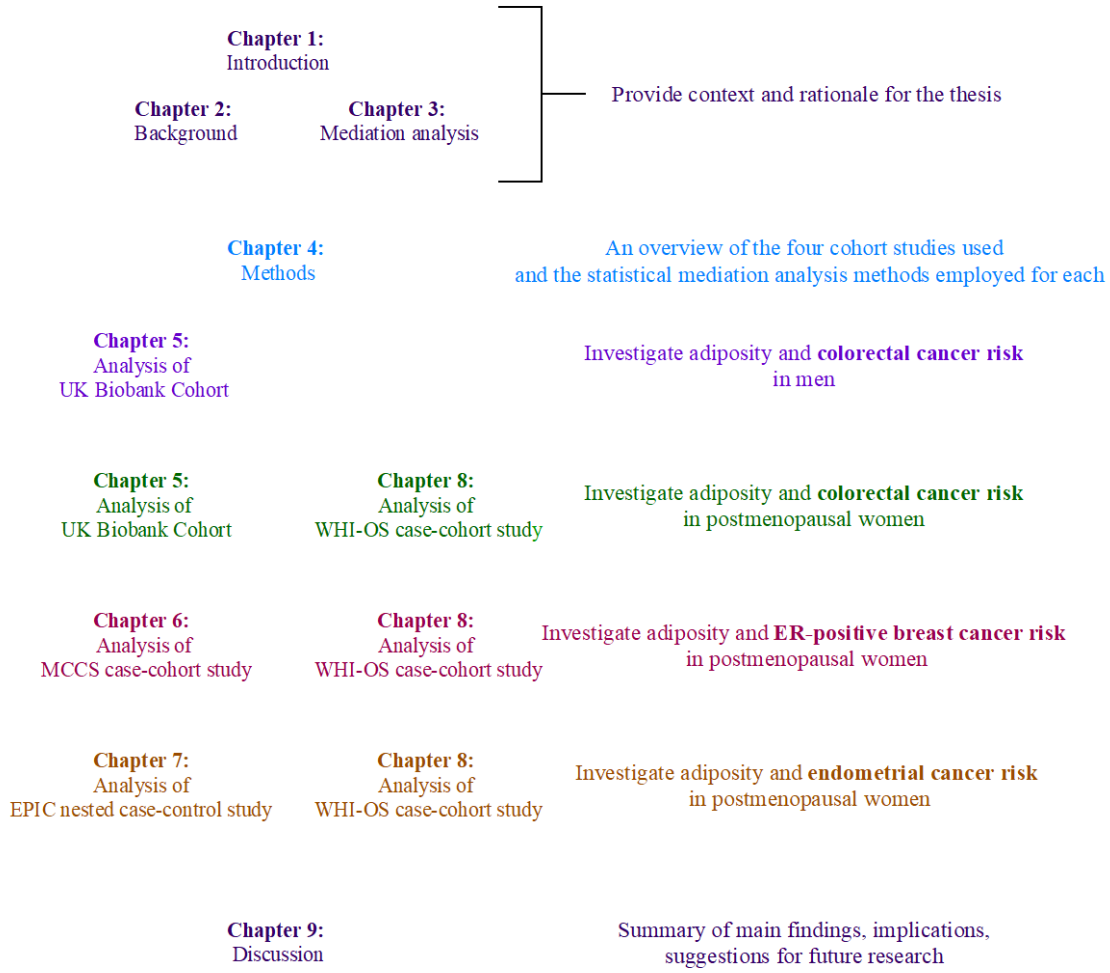
(i) disturbed adipocytokine production and chronic low-grade inflammation,

(ii) insulin resistance and hyperinsulinemia, and

(iii) alteration in levels of sex-steroid hormones.

To address the aim, causal mediation analysis approaches have been employed that allow for potential dependencies between pathways, and data from four prospective studies, namely the UK Biobank, the Melbourne Collaborative Cohort Study (MCCS), the European Prospective Investigation into Cancer and Nutrition (EPIC), and Women's Health Initiative Observational Study (WHI-OS) have been analysed. A schematic outline of the thesis is provided in Figure 1-1.

Figure 1-1 Schematic outline of thesis



Abbreviations WHI-OS Women's Health Initiative Observational Study; MCCS Melbourne Collaborative Cohort Study; EPIC European Prospective Investigation into Cancder and Nutrition; ER estrogen receptor

1.4 Significance of the thesis

Excess adiposity looms as a major cause of cancer, because its prevalence is increasing globally. Understanding the mechanisms by which adiposity affects the risk of cancer could provide important information for effective cancer prevention strategies. This thesis contributes to this knowledge by quantifying the importance of three biological pathways in explaining the effect of adiposity on the risk of colorectal, ER-positive postmenopausal breast, and endometrial cancer. Ultimately, this research will add to the body of evidence that could inform randomised controlled trials of current and novel agents that target these pathways.

Chapter 2: Background

2.1 Adiposity

The World Health Organisation (WHO) defines overweight and obesity, as "abnormal or excessive fat (adipose tissue) accumulation that may impair health" (22). Excess adiposity is a recognised risk factor for a number of chronic conditions, including diabetes, cardiovascular disease, and several types of cancer (23).

Assessment of adiposity

There are a number of measures for adiposity that are commonly used in clinical practice and epidemiological studies, including body mass index (BMI), waist circumference, and waist-hip ratio.

The most commonly used measure for adiposity is BMI, which is the weight in kilograms divided by the square of height in meters (kg/m^2). The WHO suggested cut-off values for BMI to classify adults with underweight, normal weight, overweight, or obesity are shown in Table 2-1. Measuring BMI is non-expensive and simple and has high precision and accuracy (24). However, BMI has several important limitations. It is not a suitable measure for the amount of body fat in older age groups (>65 years of age). Current BMI cut-offs used to define overweight and obesity are not appropriate for all ethnic groups. Compared with men, women have a higher percentage of body fat for the same BMI. Despite being correlated with absolute body fat and body fat percentage, BMI does not distinguish between lean and fat mass. It also is not a suitable measure for abdominal fat or the amount of visceral fat mass, which is more metabolically active and plays a more important role in disease risk than subcutaneous fat mass (24).

Waist circumference is a crude estimate of central fat and a stronger predictor of diseases related to excess adiposity than BMI but does not distinguish between visceral and subcutaneous fat mass (13, 24). Waist circumference is measured at the level of natural waistline or umbilicus. Since these anatomic locations can be difficult to identify, especially in individuals with overweight or obesity, measuring waist circumference is more prone to error compared with BMI (24). The suggested sex-specific WHO cut-off values for waist circumference are shown in Table 2-1 (25). For Asians, these thresholds do not perform equally well at identifying individuals with increased risk of disease, likely because, for the same waist circumference, Asians have a higher percentage of body fat than Europeans (24, 26).

Waist-hip ratio, which is the ratio of waist to hip circumference (measured at the widest circumference of the buttocks) is another commonly used measure for abdominal adiposity (24). Similar to waist circumference, it does not differentiate between visceral and subcutaneous fat mass. It is also measured less accurately and precisely than BMI or waist circumference (24). As a single measure of abdominal adiposity, waist circumference is regarded as superior to waist-hip ratio (24). The WHO suggested thresholds for waist-hip ratio are shown in Table 2-1 (25).

Table 2-1 The World Health Organisation suggested thresholds for body mass index, waist circumference, and waist-hip ratio for adults

	Body Mass Index (kg/m^2)	Waist circumference (cm*)	Waist-hip ratio
Underweight	<18.5	-	-
Normal weight	≥18.5 – <25	≤80 Female ≤94 Male	-
Overweight (at increased risk of metabolic complications)	≥25 – <30	>80 – ≤88 Female >94 – ≤102 Male	-
Obese (at substantially increased risk of metabolic complications)	≥30	>88 Female >102 Male	≥0.85 Female ≥0.90 Male

* The International Diabetes Federation (27) uses waist circumference ≥80 cm in women, ≥90 cm in Asian men, and ≥94 cm in Caucasian men as the cut-off points in its definition of metabolic syndrome.

Body mass index, waist circumference, and waist-hip ratio are highly correlated (Table 2-2). It is, therefore, not clear whether these measures really quantify different aspects of adiposity or which measure performs the best.

Table 2-2 Pearson's correlation coefficients between measured body mass index, waist circumference, and waist-hip ratio for participants in two large prospective cohort studies also used in this thesis, the UK Biobank and the Melbourne Collaborative Cohort Study

		BMI & waist circumference	BMI & waist-hip ratio	Waist circumference & waist-hip ratio
UK Biobank	Men n=227,423	0.88	0.59	0.79
	Women n=271,860	0.88	0.56	0.74
MCCS	Men n=17,024	0.82	0.55	0.76
	Women n= 24,444	0.84	0.44	0.75

Abbreviations BMI body mass index; MCCS Melbourne Collaborative Cohort Study

Other measures for adiposity include, but are not limited to, hip circumference, waist circumference-to-height ratio, sagittal abdominal diameter, and waist-to-thigh ratio (1). These do not appear to perform better than the previously discussed measures and are not widely used (1).

Body composition and fat distribution can also be measured using several other methods such as skinfold thickness, bioelectrical impedance (BIA), ultrasound, dual-energy X-ray absorptiometry (DXA), computed tomography (CT), and magnetic resonance imagining (MRI) (1). Compared with the previously discussed measures, these are also less commonly used, especially in large epidemiologic studies. Some (e.g. skinfold thickness, BIA, ultrasound) are less reproducible or rely on technical skills while others (e.g. DXA, CT, MRI) are expensive or difficult to access (1).

Causes of excess adiposity

Positive energy balance, which is higher energy intake (through food and drink consumption) than expenditure (through physical activity and body function), is

recognised as the main driver of excess adiposity (1, 24). At the individual level, hormonal imbalances, gut microbiota, genetics, epigenetics and maternal programming, psychosocial factors, physical activity and sedentary behaviour can influence energy balance (1, 24). Factors beyond the control of an individual, such as the built environment, affordability and accessibility of healthy diet, and nutrient composition of available foods and drinks also impact individual-level behaviours and energy balance (1, 24).

From the multiplicity of these interacting factors, it is evident that overweight and obesity are complex health conditions and both individual-level interventions and societal changes are needed to address them (1).

Trends and prevalence of excess adiposity

Adults (18 years old and older)

Globally, the prevalence of adults living with obesity (BMI >30 kg/m^2) has increased from 3.2% (95% Credible Interval (CrI)[1], 2.4% – 4.1%) in 1975 to 10.8% (95%CrI, 9.7% – 12.0%) in 2014 in men and from 6.4% (95%CrI, 5.1% – 7.8%) in 1975 to 14.9% (95%CrI, 13.6% – 16.1%) in 2014 in women (28). With its prevalence surpassing 30%, the obesity epidemic has especially affected men and women in Polynesia and Micronesia, and high-income English-speaking countries, and women in southern Africa, the Middle East, and north Africa (28). It is estimated that in 2016, there were 281 million (95%CrI, 257 – 307 million) men and 390 million (95%CrI, 363 – 418 million) women with obesity in the world (4).

[1] Credible Interval is the Bayesian equivalent of a Confidence Interval. A 95% Credible Interval is interpreted as the interval that contains the true value of the parameter of interest with 95% probability. (*Marschner IC. Inference principles for biostatisticians: Chapman & Hall/CRC; 2015.*)

The prevalence of underweight (BMI <18.5 kg/m^2) continues to be high in both sexes in south Asia (>20%) and central and east Africa (>10%). In every other region, we have long since transitioned to a world in which more adults are living with obesity than with underweight (28). Between 1975 and 2016, the age-standardised mean BMI has followed an increasing trend in all regions of the world for men and women, except for women in the high-income Asia-Pacific region, where it has been stable (4). If this trend continues, by 2025, 18% of men and 21% of women will have obesity (28).

Children and adolescents (5 to 19 years old)

The prevalence of children and adolescents with obesity (with BMI for age and sex >2 standard deviations (SD) above the median of the WHO growth reference for children and adolescents) has followed an increasing trend, from 0.9% (95%CrI, 0.5% – 1.3%) in 1975 to 7.8% (95%CrI, 6.7% – 9.1%) in 2016 in boys and from 0.7% (95%CrI, 0.4% – 1.2%) in 1975 to 5.6% (95% CrI, 4.8% – 6.5%) in 2016 in girls (4). In 2016, children and adolescents in Polynesia and Micronesia and high-income English-speaking countries had the highest prevalence of obesity, and an estimated 74 million (95%CrI, 39 – 125 million) boys and 50 million (95%CrI, 24 – 89 million) girls were living with obesity worldwide (4).

The age-standardised mean BMI has been increasing for girls and boys, with an accelerated increase post-2000 in southeast Asia for boys and in east and south Asia for both sexes. Conversely, recently in some regions, the trend has plateaued, although at high levels, for boys (south-western Europe), girls (central and Andean Latin America), or both sexes (north-western Europe, high-income English speaking and Asia-Pacific countries) (4).

Despite these increases, the prevalence of moderate and severe underweight (BMI for age and sex >2 SD below the median of the WHO growth reference for children and adolescents) has not declined much between 1975 (14.8%; 95%CrI, 10.4% – 19.5% in boys and 9.2%; 95%CrI, 6.0% – 12.9% in girls) and 2016 (12.4%; 95%CrI, 10.3% – 14.5% in boys and 8.4%; 95%CrI, 6.8% – 10.1% in girls) (4). Worldwide, there were more children and adolescents living with moderate or severe underweight than with obesity in 2016, with the highest burden observed in south Asia, and central, east, and west Africa (4). However, if the post-2000 increasing trends in mean BMI continue, by 2022, this will be reversed, and there will be more children and adolescents living with obesity in the world (4).

Disparities in prevalence of excess adiposity

The relationship between socioeconomic position (SEP) and obesity prevalence depends on the socioeconomic level of the country. In high and middle-income countries, there is an inverse relationship between the two, especially for women. On the other hand, a positive association is observed between SEP and obesity prevalence in low-income countries (1). This pattern is also observed for children (1).

In some countries, differences in the prevalence of obesity are also observed by sex (e.g. in Colombia, Iran, and Russia the prevalence is much higher for women than men) or age (e.g. in the USA, Colombia, Iran, France, and New Zealand the prevalence increases with age, then decreases at older age). These differences, however, do not appear to be associated with the socioeconomic level of the country (1).

2.2 Adiposity and cancer

In 2018, the World Cancer Research Fund (WCRF)/American Institute for Cancer Research (AICR) published their third update on "Diet, Nutrition, Physical Activity and Cancer: a Global Perspective" (29). The report concluded that there was strong, convincing evidence that excess adiposity in adulthood increased the risk of seven types of cancers, including oesophagus adenocarcinoma, and colorectal, postmenopausal breast, endometrial, pancreas, liver, and kidney cancers. A viewpoint published by the working group convened by the International Agency for Research on Cancer (IARC) in 2018 also concluded that there was sufficient evidence for a causal effect of adiposity on all of the above mentioned cancers as well as for gastric cardia, gallbladder, ovary, and thyroid cancers, meningioma, and multiple myeloma (30). The WCRF/AICR report also identified the evidence for a link between adult adiposity and three of these cancers (i.e. gastric cardia, gallbladder, ovary) as well as for cancers of the mouth, pharynx and larynx and advanced prostate cancer as strong, probable (2).

The IARC working group additionally reviewed the evidence on the association between intentional weight loss and cancer risk, adiposity and cancer survival and recurrence, and adiposity in childhood, adolescence, and young adulthood (<25 years old) and cancer risk (30). Based on the sparse existing evidence, it was proposed that intentional weight loss might lower the risk of some cancers, especially breast and endometrial cancers. A large body of evidence indicated that higher BMI close to the time of cancer diagnosis was inversely associated with survival in women with breast cancer. Evidence for other cancers was limited. Finally, the report suggested that except for postmenopausal breast cancer, for which there was evidence for an inverse association with adiposity in young adulthood, existing studies demonstrated positive associations between adiposity in

childhood, adolescence, and young adulthood and several cancers also known to be linked to adiposity in adulthood (30).

The focus of this thesis is on investigating the biological pathways that underlie the established links between adiposity in adulthood and the risk of colorectal, postmenopausal breast, and endometrial cancers. Thus, the relation between adiposity and the risk of other cancers, and the associations between weight loss and cancer risk, adiposity and cancer survival, and adiposity at other age groups and cancer risk will not be discussed further.

2.2.1 Colorectal cancer

Globally, colorectal cancer is the third most commonly diagnosed cancer in men and second in women, accounting for approximately 1 million new cases in men and 823 thousand new cases in women in 2018 (31). There is considerable heterogeneity in incidence of colorectal cancer between countries, with about 70% of the new cases and 60% of the deaths happening in countries with high or very high human development index (32). However, colorectal cancer incidence and mortality are both quickly increasing in countries with medium to high human development index, especially in countries undergoing significant development. This rise demonstrates the role some of the consequences of 'Western' lifestyle, including excess adiposity, may have on colorectal cancer risk (32).

Adiposity is an established risk factor for colorectal cancer (24, 30). It was estimated that in 2012, 13% (90% confidence interval (CI), 11% – 14%) of colon and 6% (90%CI, 5% – 7%) of rectal cancer cases diagnosed in men and 8% (90%CI, 7% – 8%) of colon and 4% (90%CI, 3% – 4%) of rectal cancer cases diagnosed in women were attributable to

having overweight or obesity (BMI ≥25 kg/m^2) (3). The population attributable fractions (PAF) were highest in countries with high and very high human development index (3).

In a recent meta-analysis of studies investigating the association between adiposity and colorectal cancer, the summary risk ratio (RR) comparing highest vs lowest waist circumference categories based on 13 cohort studies were 1.46 (95%CI, 1.33 – 1.60) overall, 1.48 (95%CI, 1.30 – 1.67) for men, and 1.44 (95%CI, 1.30 – 1.60) for women (33). For BMI, the summary RRs comparing obese vs normal weight based on 41 cohort studies were 1.33 (95%CI, 1.25 – 1.42) overall, 1.47 (95%CI, 1.36 – 1.58) for men, and 1.15 (95%CI, 1.08 – 1.23) for women. The summary RRs and 95%CIs from the WCRF dose-response meta-analyses of the studies of waist circumference, BMI, and waist-hip ratio in relation to the risk of colorectal cancer are shown in Table 2-3 (24, 34). Overall, all adiposity measures were positively associated with the risk of colorectal cancer, in both men and women, although the RR for BMI-colorectal cancer association was slightly larger for men (summary RR per 5 kg/m^2 1.08; 95%CI, 1.04 – 1.11) compared with women (summary RR per 5 kg/m^2 1.05; 95%CI, 1.02 – 1.08) (34). The difference in the strength of association for men and women was less evident for waist circumference and waist-hip ratio. When stratified by cancer site, there was a positive association for both colon and rectal cancers in relation to waist circumference and BMI, with a stronger RR observed for colon cancer (Table 2-3) (34). The observed heterogeneity by factors such as sex and subsite, might in part be due to between-study heterogeneity because specific meta-analyses included different studies.

Table 2-3 Summary of the WCRF Continuous Update Project dose-response meta-analyses for the association between adiposity and risk of colorectal cancer (34)

	Adiposity measure	Increment	N studies in meta-analysis	summary RR (95% CI)	I^2 %[2]
Overall	Waist circumference	10 cm	8	1.02 (1.01– 1.03)	0
	Body Mass Index	5 kg/m^2	38	1.05 (1.03 – 1.07)	74
	Waist-hip ratio	10%	4	1.02 (1.01 – 1.04)	17
By sex					
	Waist circumference	10 cm	4	1.02 (1.00 – 1.04)	47
Men	Body Mass Index	5 kg/m^2	20	1.08 (1.04 – 1.11)	83
	Waist-hip ratio	10%	1	1.04 (1.00 – 1.08)	NA
	Waist circumference	10 cm	5	1.03 (1.02 – 1.04)	0
Women	Body Mass Index	5 kg/m^2	24	1.05 (1.02 – 1.08)	83
	Waist-hip ratio	10%	3	1.02 (1.00 – 1.04)	34
By cancer subsite					
	Waist circumference	10 cm	10	1.04 (1.02 – 1.06)	63
Colon cancer	Body Mass Index	5 kg/m^2	41	1.07 (1.05 – 1.09)	72
	Waist-hip ratio	10%	-	-	-
	Waist circumference	10 cm	6	1.02 (1.00 – 1.03)	0
Rectal cancer	Body Mass Index	5 kg/m^2	35	1.02 (1.01 – 1.04)	59
	Waist-hip ratio	10%	-	-	-

Table adapted from the World Cancer Research Fund/ American Institute for Cancer Research report (24, 34)
Abbreviations RR risk ratio; CI confidence interval; WCRF World Cancer Research Fund

The Mendelian randomisation studies of the association between adiposity and colorectal cancer risk are also supportive of an effect of BMI (35-37) and waist circumference (35) on the risk of colorectal cancer (Table 2-4). Evidence for the effect of waist-hip ratio has been less consistent (35-37). Although the point estimates from the three studies assessing waist-hip ratio were suggestive of a positive association, only the 95% CI from one study in which the association for waist-hip ratio adjusted for BMI was assessed did not include the null value (37) (Table 2-4).

In the Mendelian randomisation study where sex-specific BMI-colorectal cancer risk associations were estimated, the point estimates were suggestive of a stronger association for women compared with men, but the 95%CIs overlapped (OR per weighted genetic risk score for BMI (i.e. having larger number of (weighted) BMI-increasing alleles) 1.55; 95%CI, 1.21 – 1.97 vs 1.09; 95%CI, 0.83 – 1.42) (37). The authors suggested that this inconsistency with observational studies might be because early-life BMI, which would

[2] I^2 is a measure of heterogeneity in a meta-analysis and is the proportion of the overall variation in estimates across studies due to heterogeneity (Higgins JP, Thompson SG. Quantifying heterogeneity in a meta-analysis. Stat Med. 2002;21(11):1539-58.)

have been better captured in the Mendelian randomisation study, might play a more critical role in increasing colorectal cancer risk in women (37).

Table 2-4 Summary of the associations between adiposity measures and colorectal cancer risk estimated in Mendelian randomisation studies

	Adiposity measure	Increment in genetically predicted measure	OR (95% CI)	Ref
Overall	Body Mass Index	1 SD	1.09 (1.01 – 1.17)	(35)
		1 SD	1.39 (1.06 – 1.82)	(36)
		Per weighted allele	1.31 (1.10 – 1.57)	(37)
	Waist circumference	1 SD	1.13 (1.02 – 1.26)	(35)
	Waist-hip ratio	1 SD	1.07 (0.91 – 1.27)	(35)
		1 SD	1.29 (0.75 – 2.22)	(36)
	Waist-hip ratio adjusted for BMI	Per weighted allele	1.36 (1.08 – 1.73)	(37)
By sex				
Men	Body Mass Index	Per weighted allele	1.09 (0.83 – 1.42)	(37)
Women	Body Mass Index	Per weighted allele	1.55 (1.21 – 1.97)	(37)

Abbreviations SD standard deviation; OR odds ratio; CI confidence interval; Ref reference

2.2.2 Breast cancer

With over 2 million new cases diagnosed in 2018, breast cancer is by far the most commonly diagnosed cancer in women globally as well as in 140 countries (31, 38). The incidence of breast cancer is higher in high-income countries, due to both higher prevalence of known risk factors for breast cancer and wider use of mammography (39), while low- and middle-income countries have higher breast cancer mortality because the disease is generally diagnosed at later stages and access to treatment is limited (39). However, breast cancer incidence is rapidly rising in countries in transition, following changes in the distribution of known risk factors for the disease, including reproductive patterns, alcohol intake, physical inactivity, and adiposity (38).

The effect of adiposity on breast cancer risk is a complex one. There is convincing evidence for a positive effect of adiposity in adulthood on the risk of postmenopausal breast cancer (24, 30). Globally, 10% (90%CI, 9% – 11%) of the postmenopausal breast cancers diagnosed in 2012 were attributable to overweight and obesity (BMI ≥25 kg/m^2), with considerably higher PAFs estimated for countries with high and very high human development index (3). On the other hand, the WCRF/AICR report concluded that there

was strong, probable evidence for an inverse relationship between adult body fatness and premenopausal breast cancer, and between body fatness in young adulthood and pre- and postmenopausal breast cancer (24). The remainder of this section describes the association between adiposity in adulthood and postmenopausal breast cancer, which is the focus of this thesis.

In a meta-analysis of 12 cohort studies investigating the association between BMI and postmenopausal breast cancer, the summary RR comparing women with BMI ≥25 – <30 vs <25 kg/m^2 and women with BMI ≥30 vs <25 kg/m^2 were respectively 1.13 (95%CI, 1.09 – 1.18) and 1.20 (95%CI, 1.11 – 1.31) (40). When stratified by hormone receptor status, no association was observed for estrogen receptor (ER)-negative/progesterone receptor (PR)-negative cancers (summary RR based on cohort and case-control studies for BMI ≥30 vs <25 kg/m^2 0.98; 95%CI, 0.78 – 1.22 vs 1.39; 95%CI, 1.14 – 1.70 for ER-positive/PR-positive cancers). Also, evidence for a positive BMI-postmenopausal breast cancer association was weaker in ever users of estrogen-progestin hormones (summary RR for BMI ≥30 vs <25 kg/m^2 1.18; 95%CI, 0.98 – 1.42) compared with never users (1.42; 95%CI, 1.30 – 1.55) (40). The results from the WCRF dose-response meta-analyses of the studies of the association between waist circumference, BMI, and waist-hip ratio and postmenopausal breast cancer, shown in Table 2-5, also supported a positive association between all adiposity measures and cancer risk (24, 41). For BMI-postmenopausal cancer association, stratified analyses were performed by cancer hormone receptor status and hormone therapy use. The positive association was limited to ER-positive, PR-positive, and ER-positive/PR-positive cancers, and to never or never/former users of hormone therapy (Table 2-5).

Table 2-5 Summary of the WCRF Continuous Update Project dose-response meta-analyses for the association between adiposity and risk of postmenopausal breast cancer (41)

	Adiposity measure	Increment	N studies in meta-analysis	summary RR (95% CI)	I^2 %
Overall	Waist circumference	10 cm	11	1.11 (1.09 – 1.13)	0
	Body Mass Index	5 kg/m^2	56	1.12 (1.09 – 1.15)	74
	Waist-hip ratio	10%	18	1.10 (1.05 – 1.16)	60
By hormone receptor status					
ER+	Body Mass Index	5 kg/m^2	14	1.17 (1.09 – 1.25)	91
PR+			5	1.47 (1.36 – 1.60)	0
ER+/PR+			9	1.29 (1.19 – 1.40)	78
ER–			13	1.00 (0.95 – 1.06)	7
PR–			5	1.05 (0.93 – 1.18)	0
ER+/PR–			6	0.94 (0.87 – 1.01)	0
ER–/PR–			9	0.96 (0.87 – 1.06)	33
By hormone therapy use					
Never/former	Body Mass Index	5 kg/m^2	4	1.20 (1.15 – 1.25)	0
Never			15	1.16 (1.10 – 1.23)	72
Current			5	0.98 (0.90 – 1.06)	69

Table adapted from the World Cancer Research Fund/ American Institute for Cancer Research report (24, 41) Abbreviations ER: estrogen receptor; PR progesterone receptor; RR risk ratio; CI confidence interval WCFR World Cancer Research Fund

The Mendelian randomisation studies of postmenopausal breast cancer have shown inverse associations for BMI (42, 43) and waist-hip ratio adjusted for BMI (43). The inverse association has been observed for never and ever hormone therapy users (Table 2-6) (42). This seeming discrepancy with observational studies might be because while the instrumental variable used in Mendelian randomisation studies captures predisposition to lifelong increased BMI, weight gained in adulthood might be the main driver of the association between adiposity in adulthood and postmenopausal breast cancer risk (42, 43).

Table 2-6 Summary of the associations between adiposity measures and postmenopausal breast cancer risk estimated in Mendelian randomisation studies

	Adiposity measure	Increment in genetically predicted measure	OR (95% CI)	Ref
Overall	Body Mass Index	5 kg/m^2	0.57 (0.46 – 0.71)	(42)
		1 SD	0.69 (0.62 – 0.77)	(43)
	Waist-hip ratio adjusted for BMI	1 SD	0.88 (0.79 – 0.99)	(43)
By hormone therapy use				
Never	Body Mass Index	5 kg/m^2	0.60 (0.38 – 0.90)	(42)
Ever	Body Mass Index	5 kg/m^2	0.47 (0.29 – 0.73)	(42)

Abbreviations SD standard deviation; OR odds ratio; CI confidence interval; Ref reference

2.2.3 Endometrial cancer

In women, endometrial cancer is the sixth most commonly diagnosed cancer globally, accounting for approximately 570,000 new cases diagnosed in 2018 (31). The incidence

is highest in North America and Europe, and lower in low- and middle-income countries (39). In many parts of the world, especially in countries with historically low levels of incidence, endometrial cancer incidence is increasing, likely due to factors such as declining rates of hysterectomy, changes in reproductive patterns, and increased prevalence of excess adiposity (39, 44).

Adiposity is one of the major risk factors for endometrial cancer (24, 45). In 2012, an estimated 34% (90%CI, 32% – 36%) of newly diagnosed endometrial cancers were attributable to overweight and obesity (BMI ≥25 kg/m^2), which of all the adiposity-related cancers was the highest estimated PAF for women (3). Similar to colorectal and postmenopausal breast cancers, the PAFs were substantially higher in countries with high and very high human development index (3).

From a meta-analysis of 20 cohort studies of the association between BMI and endometrial cancer, the summary RR for BMI ≥30 – <25 vs 18.5 – 25 kg/m^2 was 1.34 (95%CI, 1.20– 1.48) and for BMI ≥30 vs 18.5 – 25 kg/m^2 was 2.54 (95%CI, 2.27 – 2.81) (46). Table 2-7 shows the results from the WCRF dose-response meta-analyses of the studies of the association between waist circumference, BMI, and waist-hip ratio and endometrial cancer (24, 45, 47). An increased risk of cancer was observed in relation to all the three adiposity measures. The positive association was observed for both pre- and postmenopausal women. In a stratified analysis by hormone therapy use, a stronger association was found in never compared with ever users (Table 2-7) (45, 47). A stronger association with BMI has also been reported for type 1 (estrogen-driven and the most common sub-type of endometrial cancer) compared with type 2 endometrial cancer (48).

Table 2-7 Summary of the WCRF Continuous Update Project dose-response meta-analyses for the association between adiposity and risk of endometrial cancer (45, 47)

	Adiposity measure	Increment	N studies in meta-analysis	RR (95% CI)	I^2 %
Overall	Waist circumference	10 cm	4	1.28 (1.16 – 1.39)	71
	Body Mass Index	5 kg/m^2	26	1.50 (1.42 – 1.59)	86
	Waist-hip ratio	10%	5	1.21 (1.13 – 1.29)	0
By menopause status					
Premenopausal	Body Mass Index	5 kg/m^2	4	1.41 (1.37 – 1.45)	0
Postmenopausal			13	1.51 (1.38 – 1.65)	92
By hormone therapy use					
Never users	Body Mass Index	5 kg/m^2	4	1.73 (1.44 – 2.08)	87
Ever users			4	1.15 (1.06 – 1.25)	0

Table adapted from the World Cancer Research Fund/ American Institute for Cancer Research report (24, 45, 47)
Abbreviations RR risk ratio; CI confidence interval; WCRF World Cancer Research Fund

Results from Mendelian randomisation studies of the association between adiposity and endometrial cancer risk are consistent with a positive effect of BMI (Table 2-8) (49, 50). No evidence for an association between genetically predicted waist-hip ratio adjusted for BMI and endometrial cancer risk has been reported (Table 2-8) (50).

Table 2-8 Summary of the associations between adiposity measures and endometrial cancer risk estimated in Mendelian randomisation studies

Adiposity measure	Increment in genetically predicted measure	OR (95% CI)	Ref
Body Mass Index	1 SD	3.86 (2.24 – 6.64)	(49)
	1 unit weighted genetic risk score	2.24 (2.01 – 2.48)	(50)
Waist-hip ratio adjusted for BMI	1 unit weighted genetic risk score	0.97 (0.63 – 1.31)	(50)

Abbreviations SD standard deviation; OR odds ratio; CI confidence interval; Ref reference

2-3 Mechanisms linking adiposity to cancer

In the 2018 IARC Handbook of Cancer Prevention on "Absence of Excess Body Fatness", the potential mechanisms underlying the link between adiposity and cancer are categorised into two broad groups, namely intracellular factors, and metabolic and endocrinological changes (1). Figure 2-1, from the IARC Handbook demonstrates these mechanisms and their interrelations. The focus of this thesis is on the latter group. For completeness, the first section that follows draws upon the IARC Handbook (1) to briefly touch on the intracellular factors.

Intracellular factors

Dysregulated cell proliferation, apoptosis, and angiogenesis, energy-sensor pathways, and factors that compromise DNA integrity are intracellular factors (i.e. factors that operate within the target cell) that might underlie the adiposity-cancer link (1).

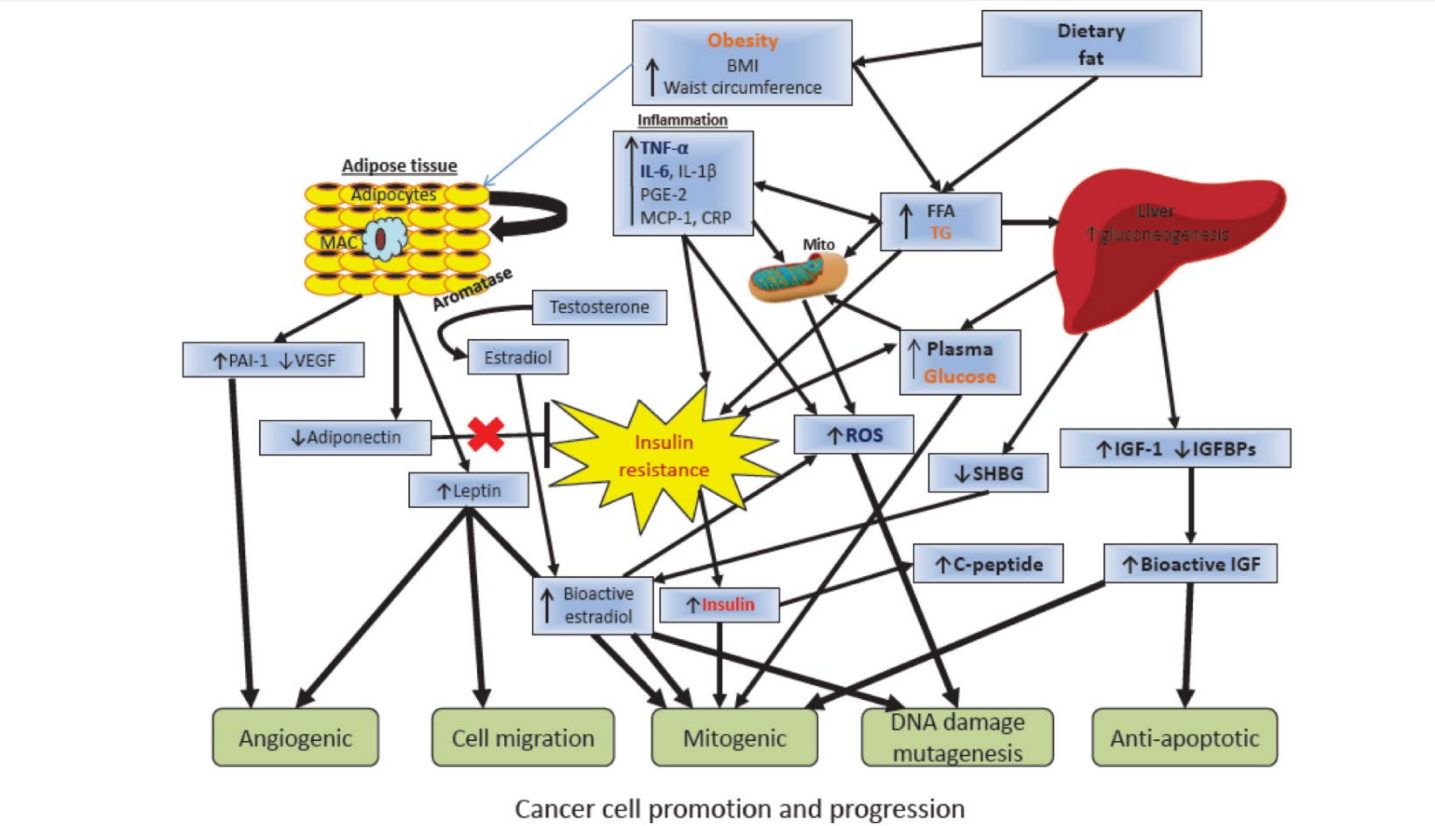
Dysregulated cell proliferation, apoptosis, and angiogenesis are established hallmarks of cancer (5). Less is known about the impact of adiposity on these processes. However, the limited evidence from human and animal studies on the effects of intentional weight loss on these suggests that adiposity might dysregulate these processes and that intentional weight loss might reverse the alterations (1).

Energy-sensor pathways control tissue growth through intracellular energy and nutrient sensors. Changes in the concentration of growth factors and hormones that occur in obesity and intentional weight loss can alter the activation of these pathways. In turn, this would influence cell proliferation, apoptosis, and angiogenesis and promote or hinder carcinogenesis. The adenosine monophosphate activated protein kinase (AMPK) – mammalian target of rapamycin (mTOR) – Protein kinase B (AKT) network (AMPK–mTOR–AKT aka the mTOR network) is one such pathway with a recognised role in the

adiposity-cancer link. Evidence for the role of other energy-sensor networks in this relation is still limited (1).

Processes that compromise DNA integrity include epigenetic mechanisms (e.g. DNA methylation), oxidative stress, impaired DNA repair function, or altered telomere length. These are emerging as mechanisms through which adiposity might promote carcinogenesis, but evidence for the mediating role of each of these processes in the adiposity-cancer link is currently limited or lacking (1).

Figure 2-1 Summary of mechanisms hypothesised to underlie the adiposity-cancer link



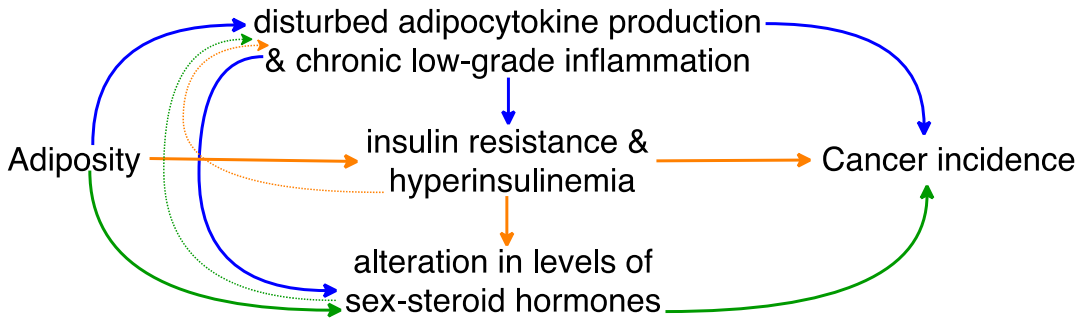
Factors denoted in bold red text are established features of the obesity–cancer connection. Factors denoted in bold blue text are emerging features.
BMI, body mass index; CRP, C-reactive protein; FFA, free fatty acids; IGF, insulin-like growth factor; IGFBP, IGF binding protein; IL, interleukin; MAC, macrophage; MCP-1, monocyte chemoattractant protein 1; Mito, mitochondria; PAI-1, plasminogen activator inhibitor 1; PGE2, prostaglandin E2; ROS, reactive oxygen species; SHBG, sex hormone-binding globulin; TG, triglycerides; TNF-α, tumour necrosis factor alpha; VEGF, vascular endothelial growth factor.
Compiled by the Working Group.

Reprinted with permission from IARC (2018). Absence of excess body fatness. IARC Handb Cancer Prev. 16:1–646. Available from: http://publications.iarc.fr/570; Chapter 4 Mechanistic and Other Relevant Data, Page 555, Copyright (2018) (1).

Metabolic and endocrinological changes

Adiposity drives three major metabolic and endocrinological alterations that might be involved in cancer development. These include disturbed adipocytokine production and chronic low-grade inflammation, insulin resistance and hyperinsulinemia, and alteration in levels of sex-steroid hormones (1, 13). Figure 2-2 is a simplified illustration of these pathways, and their postulated interrelations based on current evidence.

Figure 2-2 Diagram illustrating the three major metabolic and endocrinological pathways hypothesised to link adiposity to cancer incidence and their postulated interrelations.



The dotted thin arrows are connections for which the existing evidence is weak. These pathways were not included in the illustration of the hypothesised pathways underlying the adiposity-cancer link included in the IARC handbook (Figure 2-1 this thesis) but have been included here for completeness.

In the following sections, the link between adiposity and each of these pathways and their interrelations is discussed and an overview of the epidemiological evidence for the association between biomarkers reflecting these pathways and the risk of colorectal, postmenopausal breast, and endometrial is provided. The focus is on prospective studies (including nested case-control and case-cohort studies), since they avoid differential measurement error of exposures, biomarkers and confounders, and reverse causation. When available, results from meta-analysis or pooled analysis are used. Studies and analyses are referred to that report the biomarker-cancer association adjusted for BMI (or other measures of adiposity), since adiposity is potentially an important confounder of this association. Upstream biomarkers might also confound a biomarker-cancer association. For example, based on Figure 2-2, biomarkers of inflammation might

confound the insulin-cancer association. Therefore, when available, results adjusted for such upstream biomarkers have been reported. The IARC Handbook (1) and a comprehensive review by Nimptsch and Pischon (13) have been used as the main sources and the sections have been complemented with additional references as needed.

2.3.1 Disturbed adipocytokine production and chronic low-grade inflammation

Inflammation is a known consequence of adiposity (1). While in individuals with obesity, an intricate and not yet entirely understood network of cellular, molecular, and metabolic mechanisms is involved in causing inflammation, the roles of the following interrelated mechanisms are well established (1).

First, in the obese state, the white adipose tissue is infiltrated by macrophages, characterised by macrophages surrounding dead adipocytes known as crown-like structures, and other immune cells. The interaction between macrophages and adipocytes increases the production of pro-inflammatory cytokines, including interleukin-1β (IL-1β), IL-6, and tumour necrosis factor-α (TNF-α), from adipocytes and macrophages (1, 6).

Second, in individuals with obesity, the increased release of free fatty acids into plasma from adipose tissue increases the production of pro-inflammatory cytokines. Also, the accumulation of excess lipid in muscle, liver, and pancreas in obesity impairs lipid processing and clearance, which promotes secretion of pro-inflammatory cytokines from these tissues (1, 7).

Third, the production of adipocytokines is disturbed in the obese state (1). The production of leptin, one of the most abundantly secreted adipocytokines with pro-inflammatory, mitogenic, antiapoptotic, and angiogenic properties, is elevated with increased adipose

storage (1, 51). Conversely, the production of adiponectin, an anti-inflammatory adipocytokine, is downregulated in the hypertrophic and hyperplastic adipose tissue of individuals with obesity (1, 52). The production of other less abundantly produced pro-inflammatory adipocytokines, including plasminogen activator inhibitor-1 (PAI-1), hepatocyte growth factor (HGF), and resistin is also increased with obesity (1).

Several other mechanisms are also emerging as potential contributors to adiposity-induced inflammation. The expression of cyclooxygenase-2 (COX-2) might be increased in obesity in response to increased concentrations of free fatty acids or pro-inflammatory cytokines, such as IL-6 and TNF-α, which in turn induces adipose tissue inflammation (1, 53). Also, there is evidence suggesting that reduced bacterial diversity in the gut microbiota in individuals with obesity might cause systemic inflammation, evident by elevated C-reactive protein (CRP, an acute-phase protein and a general biomarker for inflammation) levels in plasma and white blood cells count (1).

Inflammation promotes carcinogenesis by acting as an "enabling characteristic" for several hallmarks of cancer, including cell proliferation, cell survival, and angiogenesis (5). Also, in the obese state, pro-inflammatory cytokines (6-8), as well as increased leptin and reduced adiponectin levels (9) impair insulin signalling, leading to insulin resistance and hyperinsulinemia. (These support the arrow from "inflammation" to "hyperinsulinemia" in Figure 2-2) The role of hyperinsulinemia in carcinogenesis is discussed in section **3.2.2 Insulin resistance and hyperinsulinemia.** In individuals with obesity, pro-inflammatory cytokines and leptin enhance aromatase expression, which increases estrogen production by adipose tissue (54). In men with severe obesity, pro-inflammatory cytokines and leptin might also contribute to the suppression of the hypothalamic-pituitary-testicular (HPT axis) and decrease testosterone levels (11). There

is also evidence suggesting that the production of sex hormone binding globulin (SHBG) is downregulated by pro-inflammatory cytokines TNF-α and IL-1β and upregulated by adiponectin (12). Therefore, the adiposity-induced inflammation and reduced adiponectin levels result in decreased SHBG concentrations, which in turn increases bioavailable estrogens and testosterone. (These support the arrow from "inflammation" to "sex-steroid hormones" in Figure 2-2.) The role of alterations in sex-steroid hormone levels in cancer development is discussed in section **2.3.2 Alteration in levels of sex-steroid hormones**.

Colorectal cancer

The positive association consistently seen between inflammatory bowel diseases (ulcerative colitis and Crohn's disease) and colorectal cancer risk in observational studies (55, 56), and the inverse association between aspirin use (a nonsteroidal anti-inflammatory drug (NSAID)) and colorectal cancer risk reported in observational studies and randomised controlled trials (57) provide support for a possible role of inflammation in colorectal cancer development. Studies investigating the association between inflammatory markers and colorectal cancer risk are summarised in Table 2-9.

Of the studies investigating an association between adiponectin and colorectal cancer risk (58-61), an inverse association was observed in the studies that exclusively included men (58, 61). In a nested-case control study within the European Prospective Investigation into Cancer and Nutrition (EPIC) of men and women, an inverse association was observed for non-high-molecular-weight (non-HMW) adiponectin, but not for total or HMW adiponectin (60). A Mendelian randomisation study of adiponectin and colorectal cancer risk also did not show an association between genetically determined adiponectin levels and cancer risk (62) (Table 2-9).

In the three studies of leptin and colorectal cancer risk, a positive association was reported in two that only included women (59, 63). In a large nested case-control study of men and women within the EPIC, there was evidence for an inverse association for soluble leptin receptor (sOB-R), which modulates leptin bioavailability, but not leptin (64). Stratified analyses by sex yielded similar results for men and women (P-value for difference by sex 0.58 and 0.60 for sOB-R and leptin respectively). Evidence for an inverse association between sOB-R and colorectal cancer risk was not observed in men or women in another prospective study (61) (Table 2-9).

A weak positive association was reported between CRP and colorectal cancer risk in a meta-analysis of 13 prospective studies (65). However, in the most recent and largest Mendelian randomisation study, little evidence for an association between genetically predicted increased concentration of CRP and colorectal cancer risk was observed (66) (Table 2-9).

The associations between other pro-inflammatory cytokines and colorectal cancer have been less frequently studied. In a meta-analysis of six prospective studies there was weak evidence for a small positive association for IL-6 (65). In a study reporting on plasminogen activator inhibitor-1 (PAI-1), resistin, hepatocyte growth factor (HGF), and TNF-α no association was observed (59) (Table 2-9).

Table 2-9 Associations for colorectal cancer risk for biomarkers on disturbed adipocytokine production and chronic low-grade inflammation pathway

Analysis/ Study design	N studies/ N participants	Population	Comparison/ increment	Measure	Estimate (95% CI)	I^2	Adj adiposity measure, upstream biomarker	Ref
			Adiponectin					
Cohort (HPFS)	18,225 (179 cancer affected)	Men	Q5 vs Q1	RR	0.50 (0.26 – 0.97)	-	BMI	(58)
Case-cohort (WHI-OS)	457 cases, 834 sub-cohort	Postmenopausal Women	Q4 vs Q1	HR	0.86 (0.58 – 1.29)	-	WC	(59)
Matched nested case-control (EPIC)	1,206 cases, 1,206 controls	Men and women	Q5 vs Q1	OR	0.81 (0.60 – 1.09)	-	BMI and WC	(60)
			High molecular weight adiponectin					
Matched nested case-control (EPIC)	1,206 cases, 1,206 controls	Men and women	Q5 vs Q1	OR	1.05 (0.77 – 1.43)	-	BMI and WC	(60)
			Non-high molecular weight adiponectin					
Matched nested case-control (EPIC)	1,206 cases, 1,206 controls	Men and women	Q5 vs Q1	OR	0.39 (0.26 – 0.60)	-	BMI and WC	(60)
			Adiponectin					
Matched nested case-control (HPFS)	270 cases, 519 controls	Men	Q4 vs Q1	OR	0.61 (0.38 – 0.96)	-	BMI	(61)
Matched nested case-control (NHS)	346 cases, 686 controls	Women	Q4 vs Q1	OR	1.01 (0.69 – 1.49)	-	BMI	(61)
Mendelian randomisation	1,253 cases, 1,627 controls	Men and women	Genetically determined doubling concentration	Pooled OR	0.73 (0.40 – 1.34)	-	-	(62)

continued

Table 2-9 continued Associations for colorectal cancer risk for biomarkers on disturbed adipocytokine production and chronic low-grade inflammation pathway

Analysis/ Study design	N studies/ N participants	Population	Comparison/ increment	Measure	Estimate (95% CI)	I^2	Adj adiposity measure, upstream biomarker	Ref
			Leptin					
Nested case-control (JACC)	58 cases, 145 controls	Women	Q5 vs Q1	OR	3.94 (1.04 – 14.9)	-	BMI	(63)
Case-cohort (WHI-OS)	457 cases, 834 sub-cohort	Post-menopausal Women	Q4 vs Q1	HR	1.76 (1.11 – 2.77)	-	WC	(59)
Matched nested case-control (EPIC)	1,129 cases, 1,129 controls	Men and women	Q5 vs Q1	OR	0.68 (0.43 – 1.05)	-	BMI and WC	(60)
	547 cases, 547 controls	Men			0.73 (0.43 – 1.25)			
	582 cases, 582 controls	Women			0.92 (0.55 – 1.51)			
			Soluble leptin receptor					
Matched nested case-control (HPFS)	270 cases, 519 controls	Men	Q4 vs Q1	OR	0.77 (0.48 – 1.24)	-	BMI	(61)
Matched nested case-control (NHS)	346 cases, 686 controls	Women	Q4 vs Q1	OR	1.23 (0.77 – 1.97)	-	BMI	(61)
Matched nested case-control (EPIC)	1,129 cases, 1,129 controls	Men and women	Q5 vs Q1	OR	0.61 (0.45 – 0.83)	-	BMI and WC	(64)
	547 cases, 547 controls	Men			0.69 (0.43 – 1.13)			
	582 cases, 582 controls	Women			0.57 (0.37 – 0.87)			

continued

Table 2-9 continued Associations for colorectal cancer risk for biomarkers on disturbed adipocytokine production and chronic low-grade inflammation pathway

Analysis/ Study design	N studies/ N participants	Population	Comparison/ increment	Measure	Estimate (95% CI)	I^2	Adj adiposity measure, upstream biomarker	Ref
			CRP					
Meta-analysis	13	Men and women	One-unit increment in natural log	Summary RR	1.11 (1.01 – 1.21)	59%	All for BMI	(65)
Mendelian randomisation	30,480 cases, 22,844 controls	Men and women	Genetically determined one-unit increment in natural log	OR	1.04 (0.97 – 1.12)	-	-	(66)
			IL-6					
Meta-analysis	6	Men and women	One-unit increment in natural log	Summary RR	1.10 (0.88 – 1.36)	35%	Only three for BMI	(65)
			Other					
Case-cohort (WHI-OS)	457 cases, 834 sub-cohort	Postmenopausal Women	Q4 vs Q1	HR	**PAI** 1.43 (0.93 – 2.19) **Resistin** 1.04 (0.72 – 1.50) **HGF** 1.11 (0.76 – 1.63) **TNF-α** 0.88 (0.60 – 1.29)	-	WC	(59)

Abbreviations HPFS Health Professionals Follow-Up Study; WHI-OS Women's Health Initiative Observational study; EPIC European Prospective Investigation into Cancer and Nutrition; NHS Nurse's Health Study; N number; RR Risk Ratio; HR Hazard Ratio; OR Odds Ratio; CI confidence interval; CRP C-reactive protein; IL Interleukin; PAI plasminogen activator inhibitor; HGF hepatocyte growth factor; TNF tumour necrosis factor; adj adjusted; BMI body mass index; WC waist circumference;; Ref reference

Breast cancer

The observed inverse association between aspirin use and breast cancer risk in observational studies (57) supports a potential role of inflammation in breast carcinogenesis. Studies of the association between inflammatory markers and breast cancer risk are summarised in Table 2-10.

In a meta-analysis of seven prospective studies of adiponectin and breast cancer (pre- and postmenopausal), there was weak evidence for an inverse association (67). Also, most recently, in a case-cohort study within the WHI-OS, not included in the meta-analysis, there was weak evidence of an inverse association between adiponectin and postmenopausal breast cancer (Table 2-10) (68).

A weak positive association between CRP and breast cancer was reported in a dose-response meta-analysis of 12 prospective studies (69). In the WHI-OS, also, there was weak evidence for a positive association between CRP and postmenopausal breast cancer, but the 95%CI was wide (68). In a stratified analysis by hormone therapy, the point estimates indicated a stronger association for non-users compared with users, although the 95%CIs overlapped (68). The study also assessed associations for leptin, resistin, IL-6, TNF-α, HGF, and PAI-1. The HRs were suggestive of a positive association for leptin, IL-6, HGF, PAI-1, but the 95%CIs were wide (Table 2-10) (68).

Mendelian randomisation studies of adipocytokines and pro-inflammatory cytokines and breast cancer are rare. However, a recent Mendelian randomisation analysis of adiponectin, HGF, IL-6, leptin receptor, PAI-1, and resistin and breast cancer risk, not yet published (preprint available from bioRxiv), does not support an association between any of the assessed biomarkers on and breast cancer risk (Table 2-10) (70).

Table 2-10 Associations for breast cancer risk for biomarkers on disturbed adipocytokine production and chronic low-grade inflammation pathway

Analysis/ Study design	N studies/ N participants	Population	Comparison/ increment	Measure	Estimate (95% CI)	I^2	Adj adiposity measure, upstream biomarker	Ref
			Adiponectin					
Meta-analysis	7	Pre and postmenopausal women	highest vs lowest	Summary RR	0.88 (0.73 – 1.07)	NS	Six for BMI	(67)
Case-cohort (WHI-OS)	875 cases/ 821 sub-cohort	Postmenopausal Women	Q4 and Q3 combined vs Q1	HR	0.76 (0.55 – 1.06)	-	BMI	(68)
Mendelian randomisation	122,977 cases 105,974 controls	Pre and postmenopausal women	Genetically determined unit increase	OR	1.06 (0.94 – 1.21)	-	-	(70)
			CRP					
Meta-analysis	12	Pre and postmenopausal women	Doubling concentration	Summary RR	1.07 (1.02 – 1.12)	47%	All for BMI	(69)
	9	Postmenopausal women			1.06 (1.01 – 1.11)	32%		
Case-cohort (WHI-OS)	875 cases/ 821 sub-cohort	Postmenopausal Women	Q4 vs Q1	HR	1.24 (0.86 – 1.80)	-	BMI	(68)
	401 cases/ 434 sub-cohort	Not using HT at blood draw			1.63 (0.95 – 2.80)			
	454 cases/ 363 sub-cohort	Using HT at blood draw			0.90 (0.53 – 1.53)			
Mendelian randomisation	122,977 cases 105,974 controls	Pre and postmenopausal women	Genetically determined unit increase	OR	0.96 (0.88 – 1.05)	-	-	(70)

continued

Table 2-10 continued Associations for breast cancer risk for biomarkers on disturbed adipocytokine production and chronic low-grade inflammation pathway

Analysis/ Study design	N studies/ N participants	Population	Comparison/ increment	Measure	Estimate (95% CI)	I^2	Adj adiposity measure, upstream biomarker	Ref
			Other					
Case-cohort (WHI-OS)	875 cases/ 821 sub-cohort	Postmenopausal Women	Q4 vs Q1	HR	**Leptin** 1.39 (0.93 – 2.09) **Resistin** 0.93 (0.68 – 1.27) **IL-6** 1.20 (0.85 – 1.69) **TNF-α** 0.82 (0.59 – 1.14) **HGF** 1.20 (0.87 – 1.65) **PAI-1** 1.33 (0.96 – 1.86)	-	BMI	(68)
	401 cases/ 434 sub-cohort	Not using HT at blood draw			1.71 (1.02 – 2.89)			
	454 cases/ 363 sub-cohort	Using HT at blood draw			1.17 (0.71 – 1.93)			
Mendelian randomisation	122,977 cases 105,974 controls	Pre and postmenopausal women	Genetically determined unit increase	OR	**Leptin receptor** 1.00 (0.99 – 1.01) **Resistin** 1.04 (0.94 – 1.13) **IL-6** 1.09 (0.96 – 1.25) **HGF** 1.01 (0.93 – 1.10) **PAI-1** 0.97 (0.85 – 1.10)	-	-	(70)

Abbreviations WHI-OS Women's Health Initiative Observational study; N number; HT hormone therapy; RR Risk Ratio; HR Hazard Ratio; OR Odds Ratio; CI confidence interval; CRP C-reactive protein; IL interleukin; TNF tumour necrosis factor; HGF hepatocyte growth factor; PAI plasminogen activator inhibitor; NS not stated; adj adjusted; BMI body mass index; Ref reference

Endometrial cancer

In women with overweight and obesity, but not normal weight, an inverse association has been reported between use of aspirin and other NSAIDs and endometrial cancer risk based on observational studies (71, 72), suggesting that at least in these women, inflammation might be involved in endometrial carcinogenesis. Studies of the association between inflammatory markers and endometrial cancer risk are summarised in Table 2-11.

From a meta-analysis of four nested case-control studies, there was evidence for an inverse association between adiponectin and endometrial cancer risk (73). However, in a stratified analysis by menopause status, there was no evidence for the inverse relationship in pre- and peri-menopausal women. Also, the inverse association was not observed in women using hormone therapy at blood draw (Table 2-11).

Studies on the association between other inflammatory markers and endometrial cancer risk are limited and inconsistent. Of the two studies of leptin and endometrial cancer risk, one reported a positive association in postmenopausal women not using hormone therapy at blood draw but not for hormone therapy users (P interaction 0.003) (74). The other study did not provide evidence for an association between leptin and endometrial cancer risk in postmenopausal women not using estrogen at blood draw (Table 2-11) (75).

In a case-cohort study within the WHI-OS of postmenopausal women not using hormone therapy at blood draw, a positive association was observed for CRP but not for TNF-α or IL-6 and endometrial cancer risk (76), while data from a case-control study within the EPIC did not provide evidence for an association between CRP, IL-6, or IL-1 receptor antagonist (IL-1Ra) and endometrial cancer risk in pre- and postmenopausal women (77). However, a subsequent study in this cohort supported a positive association

between TNF-α, soluble TNF receptor-1 (sTNFR-1) and sTNFR-2 and endometrial cancer risk (Table 2-11) (78).

No Mendelian randomisation analysis of pro-inflammatory cytokines or adipocytokines and endometrial cancer risk were found.

Table 2-11 Associations for endometrial cancer risk for biomarkers on disturbed adipocytokine production and chronic low-grade inflammation pathway

Analysis/ Study design	N studies/ N participants	Population	Comparison/ increment	Measure	Estimate (95% CI)	I^2	Adj adiposity measure, upstream biomarker	Ref
			Adiponectin					
Meta-analysis	4	Pre and postmenopausal	Highest vs lowest	Summary RR	0.82 (0.61 – 1.10)	0%	All for BMI	(73)
	4	Pre and postmenopausal	Increment of 10 mg/mL	Summary RR	0.82 (0.67 – 1.00)	0%	All for BMI	
	2	Pre/peri menopausal women			1.01 (0.73 – 1.40)	2%		
	4	Postmenopausal women			0.81 (0.65 – 1.00)	2%		
	3	Not using HT at blood draw			0.62 (0.44 – 0.86)	1%		
	2	Using HT at blood draw			1.10 (0.77 – 1.57)	2%		
			Leptin					
Matched nested case-control (PLCO)	167 cases, 327 controls	Postmenopausal women	T3 vs T1	OR	1.31 (0.65 – 2.64)	-	BMI	(74)
	86 cases, 173 controls	Not using HT at blood draw			4.72 (1.15 – 19.38)		BMI and estradiol*	
	79 cases, 153 controls	Using HT at blood draw			0.70 (0.24 – 2.03)			
Matched nested case-control (B~FIT)	62 cases, 124 controls	Postmenopausal women not using estrogen at recruitment	T3 vs T1	OR	2.11 (0.69 – 6.44)	-	BMI	(75)

continued

Table 2-11 continued Associations for endometrial cancer risk for biomarkers on disturbed adipocytokine production and chronic low-grade inflammation pathway

Analysis/ Study design	N studies/ N participants	Population	Comparison/ increment	Measure	Estimate (95% CI)	I^2	Adj adiposity measure, upstream biomarker	Ref
			Other					
Case-cohort (WHI-OS)	151 cases, 301 sub-cohort	Postmenopausal women not using HT at blood draw	Q4 vs Q1	HR	**CRP** 2.29 (1.13 – 4.65) **IL-6** 1.03 (0.52 – 2.03) **TNF-α** 1.52 (0.82 – 2.82)	-	BMI	(76)
Matched case-control (EPIC)	305 cases, 574 controls	Pre and postmenopausal women not using HT at blood draw	Q4 vs Q1	OR	**CRP** 1.06 (0.61 – 1.85) **IL-6** 1.33 (0.78 – 2.30) **IL-1Ra** 1.56 (0.94 – 2.58)	-	BMI	(77)
	270 cases, 518 controls	Pre and postmenopausal women not using HT at blood draw	Q4 vs Q1	OR	**TNF-α** 1.05 (0.65 – 1.69) **sTNF-R1** 1.68 (0.99 – 2.86) **sTNFR-2** 1.53 (0.92 – 2.55)	-	BMI	(78)

Abbreviations PLCO Prostate, Lung, Colorectal, and Ovarian Cancer Screening Trial; B~FIT Breast and Bone Follow-up of the Fracture Intervention Trial; EPIC European Prospective Investigation into Cancer and Nutrition; WHI-OS Women's Health Initiative Observational study; N number; HT hormone therapy; RR Risk Ratio; HR Hazard Ratio; OR Odds Ratio; CI confidence interval; CRP C-reactive protein; IL interleukin; IL-1Ra interleukin 1 receptor antagonist; TNF tumour necrosis factor; sTNF-R soluble tumour necrosis factor receptor; adj adjusted; BMI body mass index; Ref reference

*Based on the diagram in Figure 2-2 and the reviewed literature, it is unlikely that estradiol would be upstream to leptin and confound the association between leptin and endometrial cancer risk. It is more likely for estradiol to be a mediator between leptin and endometrial cancer risk. However, in stratified analysis by hormone therapy use, the paper only presented the results adjusted for estradiol levels (74).

3.2.2 Insulin resistance and hyperinsulinemia

Individuals with obesity are more likely to have insulin resistance, a condition in which sensitivity to insulin is impaired in adipose tissue, skeletal muscles, and liver (8). Insulin resistance leads to reductions in the uptake of blood glucose and the use of fatty acids by adipocytes and muscle cells, and failure in suppressing gluconeogenesis but stimulation of fatty acid production in the liver (8). Several mechanisms can drive insulin resistance in the obese state.

Excess lipid storage in adipose tissue, ectopic lipid accumulation in muscle and liver tissues, and lipotoxicity in individuals with obesity can lead to insulin resistance (8, 79).

Additionally, in obesity, cellular responses to chronic elevations of nutrients, such as activation of the endoplasmic reticulum stress response and reduced mitochondrial activity can contribute to insulin resistance (79).

The increased pro-inflammatory cytokine levels and disturbed production of leptin and adiponectin (discussed in **2.3.1 Disturbed adipocytokine production and chronic low-grade inflammation** and illustrated in Figure 2-2) can also inhibit insulin signalling and impair insulin sensitivity in obesity (8, 79). However, evidence indicates that inflammation is not necessary for insulin resistance (79). Also, more recently, a study in mice suggested that insulin resistance might itself be sufficient to induce adipose tissue inflammation (80). Although this is only emerging evidence, it highlights the complexities of the biological systems and their interrelations and interactions.

Cancer risk can be affected by elevated blood levels of insulin (i.e. hyperinsulinemia), which happens as a consequence of insulin resistance, through the mitogenic and anti-apoptotic properties of insulin (1, 13, 14). It is also hypothesised that increased insulin

levels might influence cancer risk by downregulating the production of IGF binding proteins, thereby increasing IGF-1 and bioavailable IGF-1 levels. IGF-1 has longer-term growth effects than insulin (14, 15). Insulin and IGF-1 might increase aromatase activation, thus estrogen levels, and induce activation of the estrogen receptor (ER) (10). There is also some evidence, though inconclusive, that hyperinsulinemia might downregulate SHBG production and increase bioavailable estrogens (12). (These support the arrow from "hyperinsulinemia" to "sex-steroid hormones" in Figure 2-2.) See **2.3.2 Alteration in levels of sex-steroid hormones** for a discussion on changes in sex-steroid hormone levels and cancer risk.

A large body of epidemiological evidence supports an association between increased circulating IGF-1 levels and increased risk of several types of cancer, including colorectal and breast cancers (1). However, the relationship between adiposity and IGF-1 and its binding proteins, especially based on data from epidemiological studies, is uncertain, which, in part, is because of current difficulties in measuring bioavailable IGF-1 in human samples. Some evidence suggests that there is likely a non-monotonic association between adiposity and IGF-1, with IGF-1 levels increasing with increased BMI up to ~ BMI 25 to 28 kg/m^2 , then declining with further increases in BMI (1). Therefore, IGF-1 and its binding proteins are not discussed further as biomarkers in this thesis, because the interest here is in investigating the mediating role of biomarkers in the adiposity-cancer link, where increased adiposity changes biomarker levels, and that change influences the risk of cancer.

Colorectal cancer

In an umbrella review of meta-analyses of observational studies, robust evidence for a positive association between Type 2 diabetes mellitus and colorectal cancer risk was

found (81), suggesting that hyperinsulinemia, commonly associated with diabetes, might be involved in colorectal carcinogenesis. Also, evidence from laboratory models suggests that colon cells are sensitive to insulin and that insulin has growth-promoting effects on these cells (1). Studies of the association between fasting insulin, C-peptide (a marker for insulin production), and glycosylated haemoglobin (HbA1c; a marker for glycaemic control and a correlate of insulin resistance (82)) and colorectal cancer risk are summarised in Table 2-12.

The two studies investigating associations between fasting insulin and colorectal cancer risk adjusted for an adiposity measure did not show evidence for an association (83, 84). A positive association between C-peptide and colorectal cancer was reported in three studies (Table 2-12) (85-87).

In a case-control study nested within the EPIC, there was suggestion for a weak positive association between HbA1c and colorectal cancer risk in individuals with and without diabetes (88). However, no evidence for an association was observed in a cohort study of women and men without diabetes (89). In a Mendelian randomisation analysis, also, no evidence for an association between HbA1c and colorectal cancer risk was reported (Table 2-12) (35).

Table 2-12 Associations for colorectal cancer risk for biomarkers on insulin resistance and hyperinsulinemia pathway

Analysis/ Study design	N studies/ N participants	Population	Comparison/ increment	Measure	Estimate (95% CI)	I^2	Adj adiposity measure, upstream biomarker	Ref
			Fasting insulin					
Case-cohort (WHI-OS)	438 cases, 809 sub-cohort	Postmenopausal Women without diabetes	Q4 vs Q1	HR	1.42 (0.91 – 2.23)	-	WC	(90)
Case-cohort (ATBC)	134 cases, 399 sub-cohort	Men with and without diabetes	Q4 vs Q1	HR	1.74 (0.74 – 4.07)	-	BMI	(84)
			C-peptide					
Matched nested case-control (EPIC)	1078 cases, 1078 controls	Men and women with and without diabetes	Q5 vs Q1	OR	1.37 (1.00 – 1.88)	-	BMI	(83)
Matched nested case-control (PSH)	176 cases, 294 controls	Men without diabetes	Q5 vs Q1	OR	2.7 (1.2 – 6.2)	-	BMI	(86)
Matched nested case-control (JPHC)	196 cases, 392 controls	Men with and without diabetes	Q4 vs Q1	OR	3.2 (1.4 – 7.6)	-	BMI	(87)
	179 cases, 358 controls	Women with and without diabetes			0.78 (0.38 – 1.6)			

continued

Table 2-12 continued Associations for colorectal cancer risk for biomarkers on insulin resistance and hyperinsulinemia pathway

Analysis/ Study design	N studies/ N participants	Population	Comparison/ increment	Measure	Estimate (95% CI)	I^2	Adj adiposity measure, upstream biomarker	Ref
Glycosylated haemoglobin								
Matched nested case-control (EPIC)	1026 cases, 1026 controls	Men and women with and without diabetes	Q5 vs Q1	OR	1.30 (0.96 – 1.77)	-	WHR	(88)
	561 cases, 561 controls	Men with and without diabetes			1.25 (0.83 – 1.86)			
	465 cases, 465 controls	Women with and without diabetes			1.40 (0.87 – 2.24)			
	832 cases, 832 controls	Men and women without diabetes			1.15 (0.85 – 1.57)			
Cohort study (ARIC)	6377 (129 cancer affected) 5650 (139 cancer affected)	Men without diabetes women without diabetes	≥5.7% vs 0.5%-5.6%	HR	1.04 (0.67 – 1.60) 0.99 (0.62 – 1.57)	-	BMI and WC	(89)
Mendelian randomisation	26,397 cases, 41,481 controls	Men and women with and without diabetes	Genetically determined 1 standard deviation	OR	1.02 (0·85 – 1·22)	-	-	(35)

Abbreviations WHI-OS Women's Health Initiative Observational study; ATBC Alpha-Tocopherol, Beta-Carotene Cancer Prevention; EPIC European Prospective Investigation into Cancer and Nutrition; PSH Physicians' Health Study; JPHC Japan Public Health Centre-based Prospective Study; ARIC Atherosclerosis Risk in Communities; N number; HR Hazard Ratio; OR Odds Ratio; CI confidence interval; adj adjusted; BMI body mass index; WHR waist-hip ratio; Ref reference

Breast cancer

The umbrella review of Type 2 diabetes mellitus and cancer reported a positive association with breast cancer risk (81), suggesting that hyperinsulinemia might be involved in breast cancer development. Studies of the association between fasting insulin, C-peptide, and HbA1c and breast cancer risk are summarised in Table 2-13.

In a meta-analysis of six prospective studies (five adjusted for BMI) no evidence for an association was found between fasting insulin and breast cancer (91). Heterogeneity was high in the included studies (I^2 73%), which might have been because a variety of tests had been used in the studies to measure plasma insulin, and because the two studies that mainly included premenopausal women showed inverse associations (91). The summary RR based on studies of postmenopausal women was suggestive of a stronger positive association, although the 95%CI was still wide (Table 2-13) (91). In a Mendelian randomisation study, a positive association was observed between fasting insulin and breast cancer, but there was little evidence for an association in analysis limited to premenopausal women (43). Moreover, in this analysis, the observed association was limited to ER-positive (pre- and postmenopausal) breast cancers (Table 2-13) (43).

For C-peptide and breast cancer, a meta-analysis of five prospective studies (four adjusted for BMI) yielded no evidence for an association (91). However, similar to fasting insulin, in an analysis limited to studies in postmenopausal women, the point estimate for summary RR was stronger, although the 95%CI around the estimate was wide (Table 2-13) (91).

In a large cohort study, no association between HbA1c and breast cancer was observed in individuals with and without diabetes but there was evidence for an inverse association in postmenopausal women (92). This inverse association was not observed in another

cohort study of women without diabetes, although the point estimate was again suggestive of an inverse association (Table 2-13) (89).

Table 2-13 Associations for breast cancer risk for biomarkers on insulin resistance and hyperinsulinemia pathway

Analysis/ Study design	N studies/ N participants	Population	Comparison/ increment	Measure	Estimate (95% CI)	I^2	Adj adiposity measure, upstream biomarker	Ref
			Fasting insulin					
Meta-analysis	6	Pre and postmenopausal women with and without diabetes	Q5 vs Q1	Summary RR	1.08 (0.66 – 1.78)	73%	5 for BMI	(91)
	2	Premenopausal women			0.89 (0.01 – 130.7)	87%	NS	
	3	Postmenopausal women			1.59 (0.92 – 2.74)	0%	NS	
Mendelian randomisation	182,306 participants (number affected with cancer not stated)	Pre and postmenopausal women with and without diabetes	Genetically determined 1 standard deviation	OR	1.71 (1.26 – 2.31)	-	-	(43)
		Premenopausal women			1.43 (0.74 – 2.78)		-	
		Postmenopausal women			2.07 (1.32 – 3.23)			
		ER positive cancer			2.06 (1.47 – 2.89)			
		ER-negative cancer			0.85 (0.49 – 1.47)			
			C-peptide					
Meta-analysis	5	Pre and postmenopausal women with and without diabetes	Q5 vs Q1	Summary RR	1.04 (0.77 – 1.41)	0%	4 for BMI -	(91)
	2	Premenopausal women			0.86 (0.54 – 1.38)	0%	NS	
	3	Postmenopausal women			1.26 (0.80 – 1.96)	0%	NS	

continued

Table 2-13 continued Associations for breast cancer risk for biomarkers on insulin resistance and hyperinsulinemia pathway

Analysis/ Study design	N studies/ N participants	Population	Comparison/ increment	Measure	Estimate (95% CI)	I^2	Adj adiposity measure, upstream biomarker	Ref
			Glycosylated haemoglobin					
Cohort (WHS)	27,110 (790 affected with cancer)	Pre and postmenopausal women with and without diabetes	Q5 vs Q1	HR	0.87 (0.69 – 1.10)	-	BMI	(92)
	7,442 (179 affected with cancer)	Premenopausal women			1.08 (0.65 – 1.79)			
	14,786 (499 affected with cancer)	Postmenopausal women			0.73 (0.54 – 0.98)			
	Total number not stated (136 affected with cancer)	Never HT users			0.53 (0.30 – 0.93)			
	Total number not stated (59 affected with cancer)	Past HT users			0.51 (0.21 – 1.26)			
	Total number not stated (303 affected with cancer)	HT users at blood draw			0.88 (0.60 – 1.30)			
Cohort study (ARIC)	5650 (715 cancer affected)	Postmenopausal women without diabetes	≥5.7% vs 0.5%-5.6%	HR	0.88 (0.67 – 1.15)	-	BMI	(89)

Abbreviations WHS Women's Health Study; ARIC Atherosclerosis Risk in Communities; RR risk ratio; OR odds ratio; HR hazard ratio; CI confidence interval; adj adjusted; BMI body mass index; NS not stated

Endometrial cancer

Similar to colorectal and breast cancer, a positive association between Type 2 diabetes mellitus and endometrial cancer risk reported in an umbrella review of observational studies (81) is supportive of a role of hyperinsulinemia in endometrial cancer development. Studies of the association between fasting insulin and C-peptide and endometrial cancer risk are summarised in Table 2-14.

No evidence for an association between fasting insulin and endometrial cancer risk in postmenopausal women was observed in a case-cohort study within the WHI-OS, but in a stratified analysis by hormone therapy use, there was evidence for a positive association in non-users (90). A positive effect of increased fasting insulin on increased endometrial cancer risk was also reported in a Mendelian randomisation study (Table 2-14) (49).

In a case-control study nested within three cohorts, a positive association between C-peptide and endometrial cancer risk was observed (93). Data from a nested case-control study within the EPIC were also suggestive of a positive C-peptide-endometrial cancer risk association (Table 2-14) (94).

Table 2-14 Associations for endometrial cancer risk for biomarkers on insulin resistance and hyperinsulinemia pathway

Analysis/ Study design	N studies/ N participants	Population	Comparison/ increment	Measure	Estimate (95% CI)	I^2	Adj adiposity measure, upstream biomarker	Ref
Fasting insulin								
Case-cohort (WHI-OS)	250 cases, 465 sub-cohort	Postmenopausal women without diabetes	Q4 vs Q1	HR	1.54 (0.88 – 2.67)	-	BMI	(90)
	139 cases, 288 sub-cohort	Postmenopausal women without diabetes not using HT at blood draw			2.44 (1.09 – 5.45)			
	111 cases, 177 sub-cohort	Postmenopausal women without diabetes using HT at blood draw			1.01 (0.37 – 2.79)			
Mendelian randomisation	7564 (number affected with cancer not stated)	Pre and postmenopausal women with and without diabetes	Genetically determined 1 standard deviation	OR	2.34 (1.06 – 5.14)	-	-	(49)
C-peptide								
Matched nested case-control study (NYUWHS, NSHDS, ORDET)	166 cases, 313 controls	Pre and postmenopausal women with and without diabetes	Q5 vs Q1	OR	4.40 (1.65 – 11.7)	-	BMI	(93)
Matched nested case-control study (EPIC)	286 cases, 555 controls	Pre and postmenopausal women with and without diabetes	Q4 vs Q1	OR	1.56 (0.94 – 2.57)	-	BMI	(94)

Abbreviations WHI-OS Women's Health Initiative Observational Study; NYUWHS New York University Women's Health Study; NSHDS Northern Sweden Health and Disease Study; ORDET Study of Hormones and Diet in the Etiology of Breast Cancer; EPIC European Prospective Investigation into Cancer and Nutrition; N number; HT hormone therapy; RR risk ratio; HR Hazard Ratio; OR odds ratio; CI confidence interval; adj adjusted; BMI body mass index

2.3.2 Alteration in levels of sex-steroid hormones

In postmenopausal women, estrogen and androgen levels are higher, while SHBG levels are lower for women with obesity than for women without obesity (16). After menopause, ovarian production of sex hormones ceases and adipose tissue plays a key role in estrogen production by expressing aromatase, an enzyme that converts androgens to estrogens. In the obese state, aromatase activity is enhanced in adipose tissue due to increased levels of pro-inflammatory cytokines and leptin (discussed in **2.3.1 Disturbed adipocytokine production and chronic low-grade inflammation** and illustrated in Figure 2-2) as well as hyperinsulinemia and increased IGF-1 levels (discussed in **3.2.2 Insulin resistance and hyperinsulinemia** and illustrated in Figure 2-2) (10), leading to increased synthesis of estrogens.

The mechanisms underlying increased androgen levels in women with obesity are not clearly understood (16). However, increased levels of free testosterone, as well as free estrogens, is, at least partially, due to decreased levels of SHBG in the obese state. The downregulation of SHBG in postmenopausal women with obesity might happen in response to increased TNF-α and IL-1β, decreased adiponectin, or hyperinsulinemia (12).

In men with obesity, SHBG production is also decreased (17), possibly through similar processes as in women. Contrary to women, in men with moderate obesity (BMI≥30 – <40 kg/m^2), total testosterone level is decreased, which is likely due to reduced SHBG levels. In severe obesity (BMI≥40 kg/m^2), total and free testosterone are reduced following the suppression of the hypothalamic-pituitary-testicular (HPT axis), which might be driven by increased estradiol, leptin, and pro-inflammatory cytokines (11).

Increased levels of estrogens can affect the risk of breast and endometrial cancers through their proliferative effects on these tissues (1). On the other hand, based on laboratory studies, it is suggested that estrogens might have a protective effect against colorectal cancer development via their interactions with estrogen receptor-β (ERβ), which is expressed in the colonic epithelium and has antiproliferative effects on colon tissue (1). Experimental data suggest that androgens might have also inhibitory effects on colorectal cancer tumour growth (18), thereby influencing cancer risk. Finally, there are a few studies suggesting that testosterone might suppress adiponectin production, which in turn could affect cancer risk (95, 96). Evidence supporting this effect is limited, but it again demonstrates the complexities of the interrelation between pathways underlying the link between adiposity and cancer risk.

Colorectal cancer

The association between sex-steroid hormones and colorectal cancer risk is complex and not clear (1). In women, observational studies have consistently shown an inverse association between oral contraceptive use (97), and hormone therapy (98-100), exogenous sources of sex-steroid hormones, and colorectal cancer risk. Early findings from the WHI Clinical Trial (WHI-CT) also showed an inverse association between estrogen plus progestogen intake and colorectal cancer risk (101, 102), although further follow-up of the trial suggested that the observed association might have been due to delay in cancer diagnosis rather than a true protective effect (103). Associations between endogenous sources of sex-steroid hormones and colorectal cancer risk have been less consistent (104), with an analysis of the WHI-OS showing an inverse association with ever pregnancy but not with other reproductive factors (105). Results of a Mendelian randomisation analyses of the association between age at menarche or age menopause and colorectal cancer risk were also essentially null (35, 106). Studies reporting the

association between sex-steroid hormones or SHBG and colorectal cancer risk are summarised in Table 2-15.

Of the studies examining circulating estradiol and colorectal cancer risk in postmenopausal women (107-109), only a nested case-control study within the WHI-CT, showed evidence for an inverse association (109). This study also found an inverse association with estrone and a positive association for SHBG (109). In the other two studies, there was no evidence for associations between estradiol, estrone, testosterone or SHBG and colorectal cancer risk (107, 108). Results from a Mendelian randomisation study also did not support a positive association between estradiol and colorectal cancer risk in women (Table 2-15) (35).

The only study of circulating hormones and colorectal cancer risk in men found an inverse association with testosterone and SHBG, and a positive association with estradiol to testosterone ratio (Table 2-15) (108).

Table 2-15 Associations for colorectal cancer risk for biomarkers on alteration in levels of sex-steroid hormones pathways

Analysis/ Study design	N studies/ N participants	Population	Comparison/ increment	Measure	Estimate (95% CI)	I^2	Adj adiposity measure, upstream biomarker	Ref
Matched nested case-control study (NYUWHS)	148 cases, 293 controls	Postmenopausal women	T3 vs T1	OR	**Estradiol** 0.8 (0.4 – 1.7) **Estrone** 1.6 (0.8 – 3.0) **SHBG** 0.8 (0.4 – 1.4)	-	BMI	(107)
Matched nested case-control study (NHS, WHS)	293 cases, 437 controls	Postmenopausal women not using HT at blood draw	Q4 vs Q1	OR	**Estradiol** 1.12 (0.62 – 2.03) **Estrone** 1.30 (0.74 – 2.26) **SHBG** 1.17 (0.63 – 2.20) **Testosterone** 1.43 (0.82 – 2.50)	-	BMI and C-peptide	(108)
Matched nested case-control study (WHI-CT)	401 cases, 802 controls	Postmenopausal women not using HT at blood draw	Q5 vs Q1	OR	**Estradiol** 0.64 (0.43 – 0.97) **Free estradiol** 0.43 (0.28 – 0.67) **Estrone** 0.50 (0.33 – 0.75) **SHBG** 2.30 (1.51 – 3.51)	-	WC	(109)
Mendelian randomisation	26,397 cases, 41,481 controls	Pre and postmenopausal women	Genetically determined 1 standard deviation	OR	**Estradiol** 1.27 (0.53 – 3.02)	-	-	(35)

continued

Table 2-15 continued Associations for colorectal cancer risk for biomarkers on alteration in levels of sex-steroid hormones pathways

Analysis/ Study design	N studies/ N participants	Population	Comparison/ increment	Measure	Estimate (95% CI)	I^2	Adj adiposity measure, upstream biomarker	Ref
Matched nested case-control study (HPFS, PHSII)	439cases, 719 controls	Men	Q4 vs Q1	OR	**Estradiol** 1.15 (0.73 – 1.81) **Estrone** 1.04 (0.68 – 1.62) **SHBG** 0.65 (0.42 – 0.99) **Testosterone** 0.62 (0.40 – 0.96) **Estradiol to Testosterone ratio** 2.63 (1.58 – 4.38)	-	BMI and C-peptide	(108)

Abbreviations NYUWHS New York University Women's Health Study; NHS Nurse's Health Study; WHS Women's Health Study; WHI-CT Women's Health Initiative Clinical Trial; HPFS Health Professional Follow Up Study; PHSII Physicians' Health Study II; N number; HT hormone therapy; HR hazard ratio; OR odds ratio; CI confidence interval; SHBG sex hormone binding globulin; adj adjusted; BMI body mass index; Ref reference

Breast cancer

Breast cancer is a hormonally dependent cancer and many of its established risk factors (e.g., age at menarche, age at menopause, hormone therapy) are indicators of exposure of breast tissue to estrogens (110). Results from a pooled analysis of the association between sex-steroid hormones or SHBG and postmenopausal breast cancer risk (111) are shown in Table 2-16.

A positive association has been reported between circulating estrogens (estradiol and estrone) and testosterone and postmenopausal breast cancer risk and an inverse association for SHBG (Table 2-16) (111).

Table 2-16 Associations for breast cancer risk for biomarkers on alteration in levels of sex-steroid hormones pathways

Analysis/ Study design	N studies/ N participants	Population	Comparison/ increment	Measure	Estimate (95% CI)	I^2	Adj adiposity measure, upstream biomarker	Ref
Pooled analysis of nine prospective studies	663 cases, 1765 controls	Postmenopausal women not using HT at blood draw	Q5 vs Q1	OR	**Estradiol** 2.00 (1.47 – 2.71) **Free estradiol** 2.58 (1.76 – 3.78) **Estrone** 2.19 (1.48 – 3.22) **SHBG** 0.66 (0.43 – 1.00) **Testosterone** 2.22 (1.59 – 3.10)	-	- *	(111)

Abbreviations N number; HT hormone therapy; HR hazard ratio; OR odds ratio; SHBG sex hormone binding globulin CI confidence interval; adj adjusted * not adjusted but text states that adjustment had no effect

Endometrial cancer

Estrogens, especially when unopposed by progesterone, can increase the risk of endometrial cancer through their proliferative effects in endometrial tissue (112). Established risk factors for endometrial cancer, including sequential oral contraceptives and estrogen only hormone therapy, support the "unopposed estrogen hypothesis" in endometrial cancer development (110, 112). Studies reporting the association between sex-steroid hormones or SHBG and endometrial cancer risk are summarised in Table 2-17.

Of the five prospective studies that investigated the association between estradiol and endometrial cancer risk in postmenopausal women (90, 113-116), a positive association between total, free, or unconjugated estradiol and cancer risk was observed in four studies (90, 114-116). Studies with a measure for estrone have consistently shown a positive association with endometrial cancer in postmenopausal women (114-116). Although the point estimated from the studies also suggested an inverse association for SHBG (113, 115, 116) and a positive association for testosterone (113, 115), the 95%CIs around the estimates were wide (Table 2-17).

Table 2-17 Associations for endometrial cancer risk for biomarkers on alteration in levels of sex-steroid hormones pathways

Analysis/ Study design	N studies/ N participants	Population	Comparison	Measure	Estimate (95% CI)	I^2	Adj adiposity measure	Ref
			Estradiol					
Case-cohort (WHI-OS)	250 cases, 465 sub-cohort	Postmenopausal women without diabetes	T3 vs T1	HR	**Estradiol** 3.16 (1.71 – 5.81)	-	BMI	(90)
Matched nested case-control study (EPIC)	192 cases, 374 controls	Postmenopausal women not using HT at blood draw	T3 vs T1	OR	**Estradiol** 1.61 (0.90 – 2.88) **Free estradiol** 1.03 (0.57 – 1.86)	-	BMI	(113)
Matched nested case-control study (WHI-OS)	313 cases, 354 controls	Postmenopausal women not using HT at blood draw	Q5 vs Q1	OR	**Estradiol** 1.41 (0.75 – 2.67) **Unconjugated estradiol** 6.19 (2.95 – 13.03) **Conjugated estradiol** 0.95 (0.51 – 1.77)	-	BMI	(114)
Matched nested case-control study (NYUWHS, NSHDS, ORDET)	124 cases, 236 controls	Postmenopausal women	Q4 vs Q1	OR	**Estradiol** 4.13 (1.76 – 9.72)		BMI	(115)
Matched nested case-control study (NYUWHS)	57 cases, 222 controls	Postmenopausal women	T3 vs T1	OR	**Estradiol** 1.8 (0.75 – 4.2) **Free estradiol** 2.8 (1.3 – 6.4)		BMI	(116)

continued

Table 2-17 continued Associations for endometrial cancer risk for biomarkers on alteration in levels of sex-steroid hormones pathways

Analysis/ Study design	N studies/ N participants	Population	Comparison/ increment	Measure	Estimate (95% CI)	I^2	Adj adiposity measure, upstream biomarker	Ref
Matched nested case-control study (EPIC)	192 cases, 374 controls	Postmenopausal women not using HT at blood draw	**Estrone** T3 vs T1	OR	2.19 (1.21 – 3.97)	-	BMI	(113)
Matched nested case-control study (WHI-OS)	313 cases, 354 controls	Postmenopausal women not using HT at blood draw	Q5 vs Q1	OR	**Estrone** 3.19 (1.69 – 6.04) **Unconjugated estrone** 2.70 (1.48 – 4.92) **Conjugated estrone** 2.98 (1.59 – 5.58)	-	BMI	(114)
Matched nested case-control study (NYUWHS, NSHDS, ORDET)	124 cases, 236 controls	Postmenopausal women	Q4 vs Q1	OR	3.67 (1.71 – 7.88)	-	BMI	(115)
Matched nested case-control study (NYUWHS)	57 cases, 222 controls	Postmenopausal women	T3 vs T1	OR	3.2 (1.3 – 7.8)	-	BMI	(116)

continued

Table 2-17 continued Associations for endometrial cancer risk for biomarkers on alteration in levels of sex-steroid hormones pathways

Analysis/ Study design	N studies/ N participants	Population	Comparison/ increment	Measure	Estimate (95% CI)	I^2	Adj adiposity measure, upstream biomarker	Ref
			SHBG					
Matched nested case-control study (EPIC)	192 cases, 374 controls	Postmenopausal women not using HT at blood draw	T3 vs T1	OR	0.77 (0.45 – 1.31)	-	BMI	(113)
Matched nested case-control study (NYUWHS, NSHDS, ORDET)	124 cases, 236 controls	Postmenopausal women	Q4 vs Q1	OR	0.46 (0.20 -1.05)	-	BMI	(115)
Matched nested case-control study (NYUWHS)	57 cases, 222 controls	Postmenopausal women	T3 vs T1	OR	0.49 (0.22 – 1.1)	-	BMI	(116)
			Testosterone					
Matched nested case-control study (EPIC)	192 cases, 374 controls	Postmenopausal women not using HT at blood draw	T3 vs T1	OR	1.44 (0.86 – 2.39)	-	BMI	(113)
Matched nested case-control study (NYUWHS, NSHDS, ORDET)	124 cases, 236 controls	Postmenopausal women	Q4 vs Q1	OR	1.74 (0.88 – 3.46)	-	BMI	(115)

Abbreviations WHI-OS Women's Health Initiative Observational Study; EPIC European Prospective Investigation into Cancer and Nutrition NYUWHS New York University Women's Health Study; NSHDS Northern Sweden Health and Disease Study; ORDET Study of Hormones and Diet in the Etiology of Breast Cancer; HT hormone therapy; HR Hazard Ratio; OR odds ratio; SHBG sex hormone binding globulin; CI confidence interval; adj adjusted; BMI body mass index

2.4 Summary

This chapter began by defining overweight and obesity (excess adiposity) and introducing measures of adiposity. Trends and prevalence of excess adiposity were described. Worldwide, the prevalence of obesity has been increasing in children, adolescents, and adults, and no region in the world has been able to reverse this trend over the past decades. The obesity epidemic, together with the multiplicity of factors contributing to it, many of which are outside of individuals' control, places excess adiposity as one of the most pressing public health issues of our time.

Adiposity is an established risk factor for several types of cancer, including colorectal, postmenopausal breast, and endometrial cancers. An overview of the epidemiological evidence for the effect of adiposity on these cancers was provided. Mechanisms linking adiposity to these cancers are uncertain, but disturbed adipocytokine production and chronic low-grade inflammation, insulin resistance and hyperinsulinemia, and alteration in levels of sex-steroid hormones are likely candidates. To mediate (explain) the adiposity-cancer relationship, these pathways must be influenced by adiposity and affect the risk of cancer. Epidemiological and biological evidence for the effect of adiposity on each pathway and the relationship between the pathways (as reflected by biomarkers representing them) and the three cancer sites were briefly reviewed.

For colorectal cancer, CRP shows a positive association based on a meta-analysis of observational studies, but this positive association was not replicated in a Mendelian randomisation study. The evidence on adiponectin and leptin is mostly inconsistent and there are few data on other inflammatory cytokines. There is sparse evidence showing a positive association for C-peptide but data for fasting insulin or HbA1c are not persuasive. Data for

endogenous hormones and colorectal cancer risk are also inconsistent and very limited, especially for men.

For postmenopausal breast cancer, CRP shows a positive association. The evidence on adiponectin is consistent with a weak inverse association. Few data exist on leptin or other inflammatory cytokines. Evidence is also suggestive of a positive association for fasting insulin and possibly C-peptide, but data on HbA1c are limited. There is persuasive evidence that endogenous sex hormones, especially estrogens, are positively associated with breast cancer risk.

For endometrial cancer, the limited evidence on adiponectin is consistent with a weak inverse association. Studies on the association between other inflammatory markers and endometrial cancer risk are sparse and inconsistent. Evidence is suggestive of a positive association for insulin and C-peptide. There is persuasive evidence that endogenous estrogens are positively associated with endometrial cancer risk.

Measures of the association between biomarkers and the three cancers were shown in Tables 2.9 to 2.17. Evidently, these measures do not provide any information about the contribution of the biomarkers (and the pathways they represent) to the adiposity-cancer relation. Proper mediation analysis approaches are needed to investigate and quantify the mediating roles of these biomarkers. The next chapter describes these approaches and reviews the epidemiological studies investigating the mediating roles of biomarkers in the effect of adiposity on the risk of colorectal, postmenopausal breast, and endometrial cancers.

Chapter 3: Mediation analysis

Decomposing the total effect of adiposity on the risk of cancer into indirect and direct effects, where the former is the effect(s) of adiposity through potential mediator(s) of interest and the latter is the effect of adiposity not through the investigated mediators may improve our understanding of the mechanistic pathways linking adiposity to cancer risk (19). Some of the traditional statistical approaches that allow such effect decomposition, within limits, have existed for over 30 years. In recent years, substantial advances have been made in mediation analysis methods, allowing estimation of the indirect and direct effects in complex settings. These methodological advancements have, in general, been rooted in the counterfactual framework of causation and relied on the causal inference literature, thus are broadly referred to as "causal mediation analysis" methods (19).

In the following sections, first, the basic setting of mediation analysis is described, and an overview of the traditional approaches to mediation analysis and their limitations is provided. Then, causal mediation analysis is introduced, and the no-confounding assumptions required for mediation analysis and estimation of the effects are explained. Mediation analysis approaches to handle multiple mediators and some issues for consideration in mediation analysis are also discussed. The textbook by VanderWeele, "Explanation in causal inference: methods for mediation and interaction" (19), one of the most prolific researchers in this field, is used as the primary reference for these sections.

An exhaustive review of the approaches to mediation analysis was beyond the scope of this thesis and several topics, such as mediation analysis with survival data or with time-dependent confounding, which have not been used in future chapters, are not covered here.

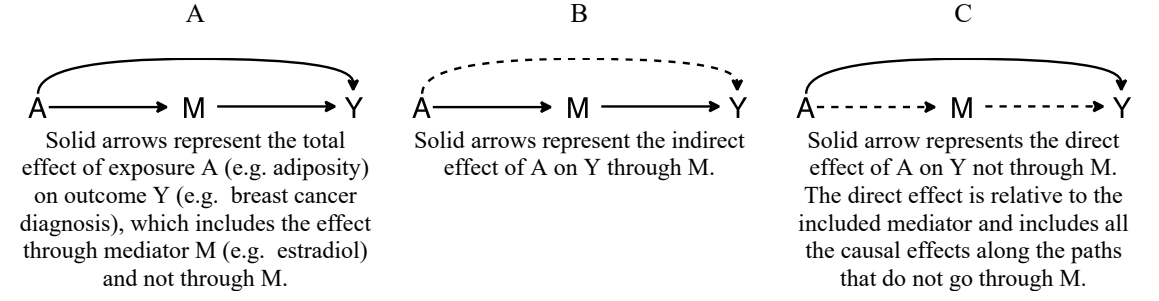
In the final section, epidemiological studies investigating the mediating effects of biomarkers on the three putative pathways linking adiposity to colorectal, postmenopausal breast, and endometrial cancer (i.e. disturbed adipocytokine production and chronic low-grade inflammation, insulin resistance and hyperinsulinemia, and alteration in levels of sex-steroid hormones; see **2-3 Mechanisms linking adiposity to cancer** for a discussion of these pathways) are reviewed.

3.1 Setting of mediation analysis

The diagram in Figure 3-1 illustrates the basic setting of mediation, where A is the exposure (e.g. adiposity), Y is the outcome (e.g. breast cancer diagnosis), and M the mediator (e.g. estradiol). The same notations will be used throughout this section and C will also be used to denote baseline confounders. Confounders are not included in Figure 3-1 for simplicity.

The total effect of A on Y is through all the solid arrows in Figure 3-1**A)**. In mediation analysis, the interest is in quantifying the part of this total effect that operates through the mediator M (i.e. the indirect effect shown by the solid arrows in Figure 3-1**B**) and not through M (i.e. direct effect shown by the solid arrow in Figure 3-1**C**) (19). Naturally, mediation is present when A changes M and that change in M subsequently affects Y. Also, the direct effect is only relative to the mediator of interest. Thus, despite what the terminology implies, it does not capture the "direct" effect of A on Y, rather the part of the effect not operating through M (19).

Figure 3-1 Mediation of the effect of an exposure on an outcome



3.2 Traditional approaches to mediation analysis

The difference method

The difference method entails a comparison between a regression model for the outcome conditional on the exposure and baseline confounders with a regression model for the outcome conditional on the exposure, baseline confounders and the mediator (19):

$$E[Y|a,c] = \beta_0 + \beta_1 a + \beta_2 c$$

$$E[Y|a,m,c] = \alpha_0 + \alpha_1 a + \alpha_2 m + \alpha_4 c$$

A material difference between β_1 and α_1 is taken as suggestion for the presence of mediation. Sometimes $\beta_1 - \alpha_1$ is taken as the estimate for the indirect effect and α_1 as the estimate for the direct effect (19).

The difference method is the most commonly used approach in papers that have assessed the potential mediating effect of biomarkers in adiposity-cancer risk, including in references (16, 59, 68, 76, 77, 83, 94, 114, 117-121), as discussed in various parts under section **2-3 Mechanisms linking adiposity to cancer**.

The product method

The product method, also famously referred to as the Barron and Kenny approach (122), involves fitting a regression model for the mediator, conditional on the exposure and baseline confounders, and a regression model for the outcome conditional on the exposure, mediator, and baseline confounders:

$$E[M|a,c] = \gamma_0 + \gamma_1 a + \gamma_2 c$$

$$E[Y|a,m,c] = \alpha_0 + \alpha_1 a + \alpha_2 m + \alpha_4 c$$

The product of γ_1 (i.e. the exposure-mediator association conditional on the confounders) and α_2 (i.e. the mediator-outcome association conditional on the exposure and confounders) is taken as the indirect effect, while again, α_1 is taken as the direct effect (19).

Limitations of traditional approaches to mediation analysis

Although the product and difference methods are intuitive and straightforward, they have several important limitations (19).

These methods do no extend to accommodate interaction between the exposure and mediator, and might, therefore, lead to biased estimation of the indirect and direct effects (19). For example, in the difference method, by ignoring interaction when one is present, but mediation is absent, one might conclude that the effect of the exposure on outcome is completely mediated, while in fact the change in the coefficient after including the mediator in the model is due to interaction and averaging the effect of exposure on outcome across the levels of the added variable (19). Even if an interaction term between the exposure and mediator is included in the regression model for the outcome, the difference and product methods do not specify how to deal with the interaction coefficient when estimating mediation (19).

Also, neither of the approaches allows for nonlinear exposure-outcome or mediator-outcome associations. As with the interaction term, even if a term for nonlinearity is included in the outcome model, it is not clear how it must be dealt with in the estimation of the mediated effect (19).

Another limitation of these methods arises when handling non-rare binary outcomes using logistic regression models. Logistic regression estimates the odds ratio, which is a "non-collapsible" measure (123). Therefore, the odds ratio estimated from a model for the outcome not including the mediator, and the odds ratio estimated from a model for the outcome

including the mediator are not directly comparable (19). The issue of non-collapsibility of the odds ratio in assessing mediation can be intuitively understood when considering the difference method, where the regression coefficient might increase in the model including the mediator because of the additional covariate (mediator) but decrease because of mediation. This would lead one to mistakenly conclude that there is no mediation since the coefficients from the models with and without the mediator are essentially the same (19). Odds ratios for rare outcomes and risk ratios, estimated from log-linear models, do not have this problem (19).

Finally, the difference and product methods do not extend to situations where there are multiple mediators affecting or interacting with each other (19).

Results from the difference and product methods coincide when the outcome is continuous or rare binary. They give an unbiased estimation of the mediated effect when, in addition, there is no exposure-mediator interaction, no nonlinear exposure-outcome or mediator-outcome associations, no measurement error, and no violation of the no-confounding assumptions required for mediation analysis (discussed in **3.4 No confounding assumptions for mediation analysis**). The rare binary outcome requirement can be relaxed if a log-linear model is used instead of a logistic regression model (19).

3.3 Causal mediation analysis

Causal mediation analysis, rooted in the counterfactual framework, allows effect decomposition even in the presence of exposure-mediator interaction and nonlinearities. More recently, extensions of this approach have become available that allow estimation of the indirect and direct effects in the presence of multiple interrelated and interacting mediators (19). These advancements are of particular interest for studies that aim to investigate the mediating role of several interrelated mechanistic pathways (e.g. inflammatory status, hyperinsulinemia, and sex-steroid hormones) in the effect of adiposity on cancer risk. Another

important contribution of the causal inference literature to the area of mediation analysis has been clarification of the no-confounding assumptions required to estimate the mediating effect. These assumptions are discussed in **3.4 No confounding assumptions for mediation analysis**.

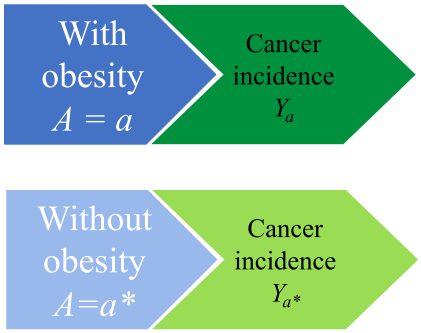
The most commonly identified effects under the causal mediation analysis approach are the average natural indirect effect (NIE), the average natural direct effect (NDE), and the average controlled direct effect (CDE) (19). Individual level counterfactual definitions of these effects (124, 125) are provided below, based on the following notations and assuming a binary exposure for simplicity.

For a binary exposure A, where a is exposed and a^* unexposed, there are two potential outcomes for an individual: Y_a, outcome when exposed, and Y_{a^*}, outcome when unexposed, and two potential mediator values: M_a, mediator value when exposed, and M_{a^*}, mediator value when unexposed (19).

Under the counterfactual framework, the individual level total effect (*te*[3]) of A on Y is $Y_a - Y_{a^*}$, which is the contrast between the outcome had the individual been exposed and the outcome had they been unexposed (Figure 3-2).

[3] Lower case italic notations (i.e. *te*, *nie*, *nde*, and *cde*) are used to denote the individual level effects.

Figure 3-2 Schematic illustration of the counterfactual definition of the total effect of adiposity, as a binary variable (with vs without obesity), on cancer incidence



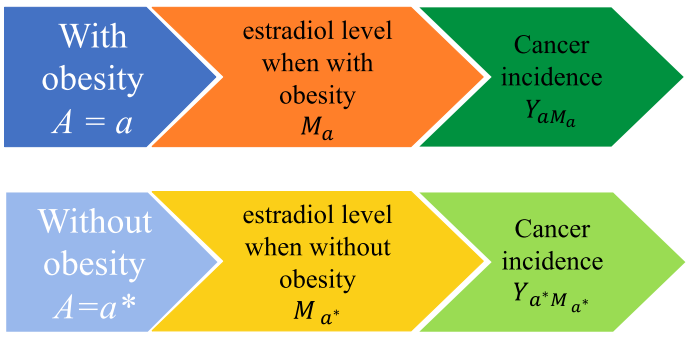
$$te = Y_a - Y_{a^*}$$

Average total effect could be written as $\mathrm{E}(te) = \mathrm{TE} = \mathrm{E}[Y_a - Y_{a^*}]$

Abbreviations *te* individual level total effect; TE average total effect

This could be equally written as $Y_{aM_a} - Y_{a^*M_{a^*}}$, which is the outcome had the individual been exposed, and their mediator level was at the level it would have naturally been when exposed and the outcome had they been unexposed, and their mediator was at the level it would have naturally been when unexposed (Figure 3-3) (19).

Figure 3-3 Schematic illustration of the counterfactual definition of the total effect of adiposity, as a binary variable (with vs without obesity), on cancer incidence



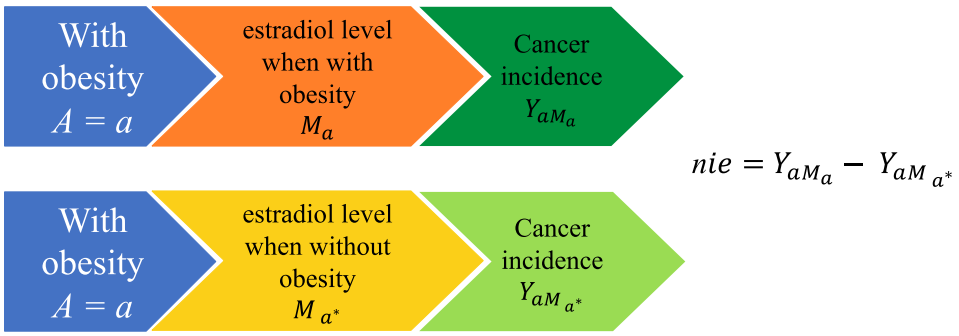
$$te = Y_{aM_a} - Y_{a^*M_{a^*}}$$

Average total effect could be written as $\mathrm{E}(te) = \mathrm{TE} = \mathrm{E}[Y_{aM_a} - Y_{a^*M_{a^*}}]$

Abbreviations *te* individual level total effect; TE average total effect

The individual level natural indirect effect (*nie*) is defined as $Y_{aM_a} - Y_{aM_{a^*}}$, which is the contrast between the outcome had the individual been exposed, and their mediator level was at the level it would have naturally been when exposed and the outcome had they been exposed, but their mediator was at the level it would have naturally been when unexposed. The *nie* would be non-zero if the mediator level when exposed were different from when unexposed (i.e. exposure changed the mediator) and that difference affected the outcome. It thus captures the effect of exposure on outcome through the mediator (Figure 3-4) (19).

Figure 3-4 Schematic illustration of the counterfactual definition of the natural indirect effect of adiposity, as a binary variable (with vs without obesity), on cancer incidence through estradiol levels



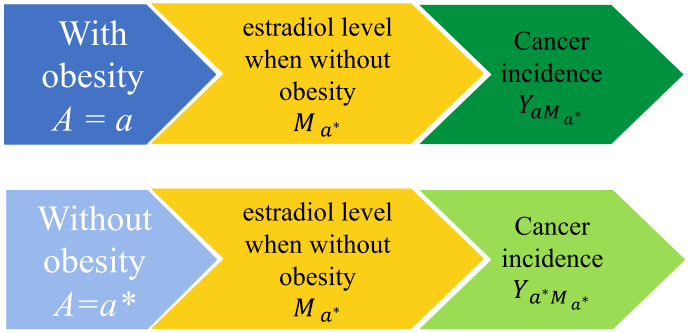
$$nie = Y_{aM_a} - Y_{aM_{a^*}}$$

Average natural indirect effect could be written as $E(nie) = \text{NIE} = E[Y_{aM_a} - Y_{aM_{a^*}}]$

Abbreviations *nie* individual level natural indirect effect; NIE average natural indirect effect

The individual level natural direct effect (*nde*) is defined as $Y_{aM_{a^*}} - Y_{a^*M_{a^*}}$, which is the contrast between outcome had the individual been exposed, but their mediator level was at the level it would have naturally been when unexposed, and outcome had they been unexposed, and their mediator was at the level it would have naturally been when unexposed. The *nde* captures the effect of exposure on outcome if it were possible to remove the influence of exposure on the mediator level (Figure 3-5) (19).

Figure 3-5 Schematic illustration of the counterfactual definition of the natural direct effect of adiposity, as a binary variable (with vs without obesity), on cancer incidence not through estradiol levels



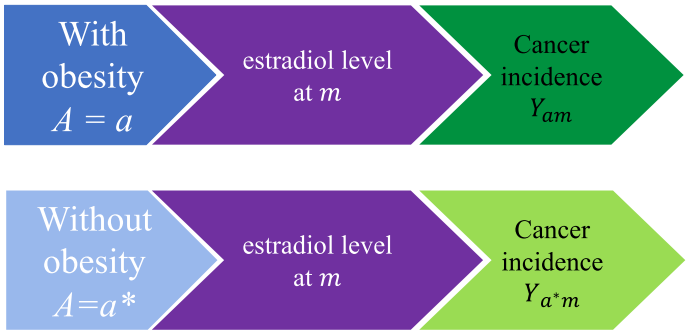
$$nde = Y_{aM_{a^*}} - Y_{a^*M_{a^*}}$$

Average natural direct effect could be written as $E(nde) = \text{NDE} = E[Y_{aM_{a^*}} - Y_{a^*M_{a^*}}]$

Abbreviations *nde* individual level natural direct effect; NDE average natural direct effect

Finally, the individual level controlled direct effect (*cde*) is $Y_{am} - Y_{a^*m}$, which is the contrast between the outcome had the individual been exposed and their mediator level was fixed at a given value m and the outcome had they been unexposed, and their mediator level was again fixed at value m. Similar to the nde, this effect captures the effect of exposure on outcome, while fixing the mediator value, thus removing the influence of exposure on the mediator. Contrary to the *nde*, for which the mediator level is fixed at a level it would have naturally taken when unexposed, for *cde* the mediator level is fixed at an arbitrary value. In the absence of exposure-mediator interaction the *nde* and *cde* coincide (Figure 3-6) (19).

Figure 3-6 Schematic illustration of the counterfactual definition of the controlled direct effect of adiposity, as a binary variable (with vs without obesity), on cancer incidence when estradiol level is fixed at level m



$$cde = Y_{am} - Y_{a^*m}$$

Average controlled direct effect could be written as $E(cde) = \text{CDE} = E[Y_{am} - Y_{a^*m}]$

Abbreviations *cde* individual level controlled direct effect; CDE average controlled direct effect

One of the most important and desirable properties of the *nie* and *nde* is that their sum is equal to the *te* (or if calculated on the multiplicative scale their product):

$$nie + nde = Y_{aM_a} - Y_{aM_{a^*}} + Y_{aM_{a^*}} - Y_{a^*M_{a^*}} = Y_{aM_a} - Y_{a^*M_{a^*}}$$

In other words, the *te* decomposes into the *nie* and *nde* (or *cde* when exposure-mediator interaction is absent). In the presence of exposure-mediator interaction, the sum of *nie* and *cde* would diverge from the *te* and difference between the *te* and *cde* cannot be interpreted as an indirect effect, thus the *nde* is preferred over the *cde* for effect decomposition (19).

3.4 No confounding assumptions for mediation analysis

Evidently, the individual level effects described above (i.e. *te*, *nie*, *nde*, and *cde*) are not observable. These effects can only be estimated on average at the population level (i.e. TE, NIE, NDE, and CDE) under a number of no-confounding assumptions (19). These assumptions illustrated in Figure 3-7, include:

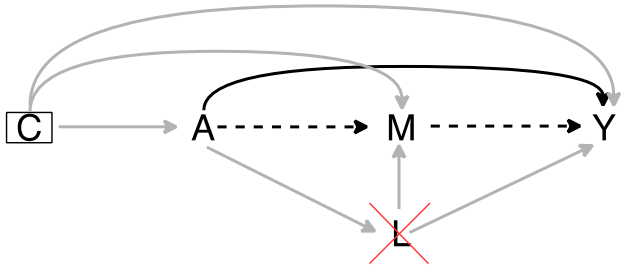
(i) no unmeasured exposure-outcome confounding (as in any observational study),

(ii) no unmeasured mediator-outcome confounding (because in mediation analysis conclusions are made about the effect of mediator on outcome),

(iii) no unmeasured exposure-mediator confounding (because in mediation analysis conclusions are made about the effect of exposure on mediator), and

(iv) no mediator-outcome confounder that is affected by the exposure (19).

Figure 3-7 Causal diagram illustrating the no-cofounding assumptions required for mediation analysis



In this diagram, *A* is the exposure, *Y* the outcome, and *M* the mediator. C depicts a vector of baseline confounders. Under the no-confounding assumptions (i) to (iii), conditional on C, there is no unmeasured confounding for A-Y, M-Y, and A-M associations. L is an M-Y confounder that itself is affected by A. Under assumption (iv), L does not exist.

In Figure 3-7, *L* is a confounder of the *M* - *Y* association that is itself affected by *A*. Handling scenarios where *L* exist is complicated because even if *L* is measured, both adjusting and not adjusting for it would introduce bias in the estimates of the NDE and NIE (19). Extensions of causal mediation analysis have been developed that allow effect decomposition when *L* exists (19). These approaches are discussed in section **3.7 Mediation analysis with multiple mediators**.

All four no-confounding assumptions need to be met for the estimates of NIE and NDE to have causal interpretations. For the CDE, only assumptions (i) and (ii) are required, because fixing the value of mediator to a specific level essentially removes any arrow influencing *M* (19).

In addition to the no-confounding assumptions described above, which are required to achieve conditional exchangeability, identifying the NIE, NDE, and CDE relies on consistency and

positivity assumptions (126). Consistency requires that, for each individual, the counterfactual quantity Y_{am} is equal to the outcome value that is observed when the individual's exposure value is $A = a,$ and the mediator value $M = m$. Positivity requires that conditional on confounders, the probability of $A = a$ or $A = a^*$ is nonzero (positive), and conditional on confounders and A, the probability of $M = m$ is nonzero for all values of M (126).

3.5 Estimation of the effects under causal mediation analysis

Continuous outcome

Based on the definitions provided in section **3.3 Causal mediation analysis** and assuming the four no-confounding assumptions are met, for a continuous outcome and comparing change in exposure from A = a to A = a*, the NIE, NDE, and CDE conditional on C=c (for some level *c* of the confounder *C*) can be empirically identified as follows (19):

$$E[nie|C = c] = NIE(c) = \sum_m E[Y|A = 1, m, c]\{P(M = m|A = a, c) - P(M = m|A = a^*, c)\}$$

$$E[nie|C = c] = NDE(c) = \sum_m \{E[Y|A = a, m, c] - E[Y|A = a^*, m, c]\}\ P(M = m|A = a^*, c)$$

$$E[cde_m|C = c] = CDE_m(c) = \ E[Y|A = a, m, c] - E[Y|A = a^*, m, c]$$

For a continuous outcome and mediator, comparing exposure $A = a$ with $A = a^*$, and allowing for exposure-mediator interaction, the following regression models can be used to estimate the above (19):

$$E[M|a, c] = \gamma_0 + \ \gamma_1 a + \ \gamma_2 c$$

$$E[Y|a, m, c] = \alpha_0 + \ \alpha_1 a + \alpha_2 m \ + \alpha_3 am + \ \alpha_4 c$$

As follows (19):

$$NIE(c) = (\alpha_2\gamma_1 + \alpha_3\gamma_1 a)(a - a^*)$$

$$NDE(c) = (\alpha_1 + \alpha_3\gamma_0 + \alpha_3\gamma_1 a^* + \alpha_3\gamma_2 c)(a - a^*)$$

$$CDE_m(c) = (\alpha_1 + \alpha_3 m)(a - a^*)$$

As explained before, to estimate the CDE, only no-confounding assumptions (i) and (ii) are required (19). For a binary mediator, a logistic regression model needs to be fit and the formulae above can easily be modified (available from reference (19)).

Binary outcome

Following from above, under the four no-confounding assumptions, for a binary outcome and considering a change in exposure from $A = a$ to $A = a^*$, the NIE, NDE, and CDE conditional on C can be identified on the odds ratio (OR) scale as follows (19):

$$OR^{NIE}(c) = \frac{P\left(Y_{aM_a} = 1 \,|c\right)/\{1 - P\left(Y_{aM_a} = 1 \,|c\right)\}}{P\left(Y_{aM_{a^*}} = 1 \,|c\right)/\{1 - P\left(Y_{aM_{a^*}} = 1 \,|c\right)\}}$$

$$OR^{NDE}(c) = \frac{P\left(Y_{aM_{a^*}} = 1 \,|c\right)/\{1 - P\left(Y_{aM_{a^*}} = 1 \,|c\right)\}}{P\left(Y_{a^*M_{a^*}} = 1 \,|c\right)/\{1 - P\left(Y_{a^*M_{a^*}} = 1 \,|c\right)\}}$$

$$OR^{CDE}(m|c) = \frac{P(Y_{am} = 1 \,|c)/\{1 - P(Y_{am} = 1 \,|c)\}}{P(Y_{a^*m} = 1 \,|c)/\{1 - P(Y_{a^*m} = 1 \,|c)\}}$$

For a rare binary outcome and continuous mediator, comparing exposure $A = a$ with $A = a^*$, and allowing for exposure-mediator interaction, the following regression models can be used to estimate the above (19):

$$E[M|a, c] = \gamma_0 + \gamma_1 a + \gamma_2 c$$

$$logit(P(Y = 1|a, m, c)) = \alpha_0 + \alpha_1 a + \alpha_2 m + \alpha_3 am + \alpha_4 c$$

As follows (19):

$$log(OR^{NIE}(c)) \cong (\alpha_2\gamma_1 + \alpha_3\gamma_1 a)(a - a^*)$$

$$log(OR^{NDE}(c)) \cong (\alpha_1 + \alpha_3\gamma_0 + \alpha_3\gamma_1 a^* + \alpha_3\gamma_2 c + \alpha_3\sigma^2)(a - a^*) + 0.5\alpha_3^2\sigma^2(a^2 - a^{*2})$$

$$log(OR^{CDE}(c)) = (\alpha_1 + \alpha_3 m)(a - a^*)$$

Again, to estimate the OR^{CDE}, only no-confounding assumptions (i) and (ii) are required and the rare outcome requirement is not needed. For the OR^{NDE}, an additional assumption for the normal distribution of M is also required. In the formula for the OR^{NDE}, σ^2 is the error variance in the regression model for M (19). As evident from the formula, this variance only influences the estimate for the OR^{NDE} when an interaction term between A and M has been included. For the OR^{NIE} and OR^{NDE}, the rare outcome assumption can be relaxed if instead of a logistic regression model a log-linear model is used (19). The above formulae can be modified for a binary mediator and are available from reference (19).

Under the rare outcome assumption, the approach described here can easily be adapted for an unmatched case-control study by fitting the mediator model for the controls, assuming that the control group is a true random sample from the source population (127). Similarly, for case-cohort data, the mediator model can be fitted for the sub-cohort group. For a matched case-control study, since the distribution of the baseline confounders in the control group might not be the same as the underlying cohort, either the estimated effects need to be reported at specific levels of the baseline confounders, or if the baseline confounders are set to their averages in the matched control group for estimation of the indirect and direct effects, the interpretation of these effects needs to be modified to reflect this. Alternatively, the effects can be estimated

with baseline confounders set to their averages in the source population if these data are available (128).

3.6 Proportion mediated

To give some intuition for the importance of a mediator in explaining an exposure-outcome association, sometimes a measure called the proportion mediated (PM) is calculated. On the risk difference scale, PM is simply (19):

$$PM = \frac{NIE}{TE} = \frac{NIE}{NIE + NDE}$$

If the effects have been estimated on the relative scale, PM can still be calculated using the equation above. For effects estimated on the OR scale, the formula will be (19):

$$PM = \frac{OR^{NDE}(OR^{NIE} - 1)}{OR^{NDE} \times OR^{NIE} - 1}$$

Alternatively, PM can be estimated on the log-odds scale, which is an additive scale.

This PM aims to capture how much the effect of exposure on outcome would be reduced in relative terms if it were possible to disconnect the link between the exposure and the mediator. However, although at times helpful, PM has several limitations. First, it is highly unstable and the confidence intervals around it are often wide. It is thus recommended to refer to the confidence interval around the NIE to comment on evidence for presence of a mediated effect instead (19). An additional problem arises when the NDE and NIE are in opposite directions, leading to PMs larger than 100%. Therefore, it is advised to restrict calculating PM to situations where NIE and NDE are in the same direction (19).

Sometimes, another measure, proportion elimination (PE) is calculated as follows on the risk difference scale:

$$PE(m) = \frac{TE - CDE(m)}{TE}$$

The effects calculated on the OR scale can be used to estimate the PE on the excess relative risk scale:

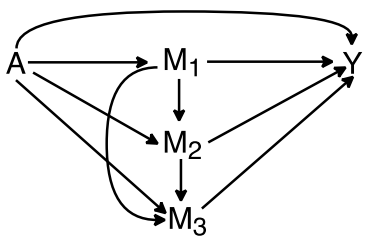
$$PE(m) = \frac{OR^{TE} - OR^{CDE}(m)}{OR^{TE} - 1}$$

This measure aims to capture how much the effect of exposure on outcome would be reduced if it were possible to intervene on the mediator and fix it at a specific level. When contrasted with the interpretation for PM (how much the effect of exposure on outcome would be reduced if it were possible to disconnect the link between the exposure and mediator), it becomes apparent that of the two measures, PE might be more relevant for policy and PM for understanding aetiology. If calculated on the same scale and in the absence of exposure-mediator interaction PM and PE will coincide (19).

3.7 Mediation analysis with multiple mediators

So far, mediation analysis with a single mediator has been discussed. When there are multiple mediators of interest, one approach would be to assess the mediated effect for each mediator separately. However, this approach would violate the fourth no-confounding assumption (i.e. no mediator-outcome confounder affected by the exposure) for some of the mediators if the mediators affect one another, as illustrated in Figure 3-8 (19).

Figure 3-8 Diagram illustrating the violation of no mediator-outcome confounder affected by the exposure in the presence of multiple mediators affecting each other



In this diagram, A is the exposure, Y the outcome, and M_1, M_2, M_3 the mediators. Here, M_1 acts a mediator-outcome confounder affected by the exposure for M_2 and M_3, and M_2 acts a mediator-outcome confounder affected by the exposure for M_3.

Even if the mediators do not affect each other but only interact with each other, assessing mediation through each mediator individually would fail (i.e. the sum [or product] of the NIEs would be larger than the TE) (19).

Extensions of causal mediation analysis are available that allow effect decomposition when several mediators affecting and interacting with each other are of interest (19). Three of these approaches, which are the most common and have been used in subsequent chapters of this thesis, are described below. The example in Figure 3-8 will be referred to throughout to illustrate the effects identified under each of these approaches. For simplicity, confounders are not included in Figure 3-8. This figure is similar to Figure 2-2, where A is adiposity, Y cancer incidence, M_1 disturbed adipocytokine production and chronic low-grade inflammation, M_2 insulin resistance and hyperinsulinemia, and M_3 alteration in levels of sex-steroid hormones.

3.7.1 Joint analysis

Using similar counterfactual definitions in to section **3.3 Causal mediation analysis**, the NIE can be identified through multiple mediators (e.g. M_1, M_2, and M_3) jointly, and the NDE or CDE operating through pathways other than M_1, M_2, and M_3 (Figure 3-9). The formulae presented in **3.5 Estimation of the effects under causal mediation analysis** can be used by replacing the single mediator M with $\boldsymbol{M}$, representing a vector of mediators of interest (i.e. $\boldsymbol{M} = (M_1, M_2, M_3)$). The four no-confounding assumptions must hold for **M** and any other [measured] mediator-outcome confounder affected by the exposure can be added to the vector of mediators to avoid violation of the fourth assumption (129). No assumption about the causal ordering of the mediators is required for identification of the effects through and not through the mediators jointly.

Figure 3-9 Diagram illustrating estimating indirect and direct effects with M representing the vector of mediators M_1, M_2, M_3

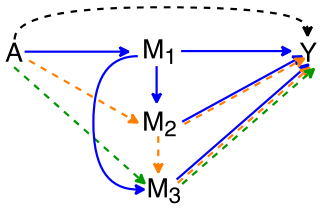
In this diagram, A is the exposure, Y the outcome, and $\boldsymbol{M}$ is the vector of mediators M_1 , M_2, M_3.

The regression models described above (**3.5 Estimation of the effects under causal mediation analysis**) can be used to estimate the NIE and NDE for $\boldsymbol{M}$ (129).

3.7.2 Sequential analysis

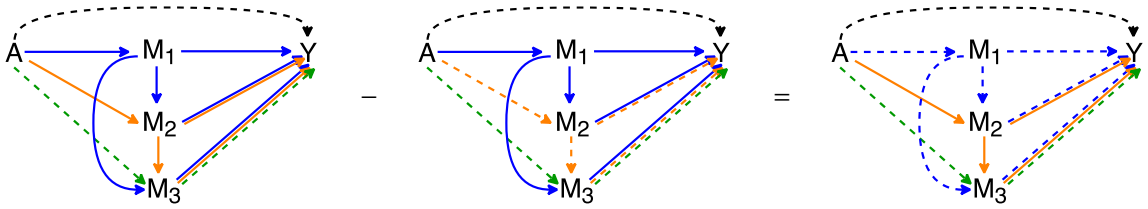
If the multiple mediators are sequential and their causal ordering is known, some progress beyond assessing the joint mediating effect can be made, where the NIEs through M_1 (the combination of $A \rightarrow M_1 \rightarrow Y$; $A \rightarrow M_1 \rightarrow M_2 \rightarrow Y$; $A \rightarrow M_1 \rightarrow M_2 \rightarrow M_3 \rightarrow Y$; $A \rightarrow M_1 \rightarrow M_3 \rightarrow Y$), through M_2, excluding (beyond) the influence of M_1 on M_2 (the combination of $A \rightarrow M_2 \rightarrow Y$; $A \rightarrow M_2 \rightarrow M_3 \rightarrow Y$), and through M_3 beyond the influence of M_1 and M_2 on M_3 ($A \rightarrow M_3 \rightarrow Y$) can be identified. The process for this sequential mediation analysis is illustrated in Figure 3-10 (129).

Figure 3-10 Diagrams illustrating the sequential mediation analysis approach



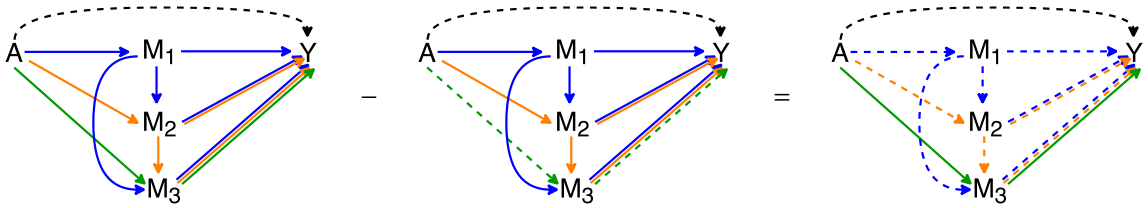
In these diagrams, A is the exposure, Y the outcome, and M_1 , M_2, M_3 the mediators.
Under the sequential mediation analysis approach, first the natural indirect effect (NIE) through M_1 can be estimated. This NIE also captures the NIE through M_1 operating to the influences of M_1 on M_2 and M_3 (the solid blue arrows). The NIE through M_1 in isolation (i.e. not through influencing M_2 and M_3) cannot be identified.

A B C



Then the joint NIE through M_1 and M_2 (solid blue and orange arrows in A) can be estimated. The difference between this (A) and the NIE through M_1 (B) gives the NIE through M_2 excluding the influence of M_1on M_2 (C). Again, the NIE through M_2 excluding its influence on M_3cannot be identified.

A B C



Finally, the joint NIE through M_1, M_2, M_3 (solid blue, orange, and green arrows in A) can be estimated. The difference between this (A) and the NIE through $M_1$and M_2 (B) gives the NIE through M_3 excluding the influences of $M_1$and M_2 (C).

The no-confounding assumptions (i), (ii), and (iii) must hold for each mediator and any [measured] exposure-induced mediator-outcome confounder cancer be included as an additional mediator in the analysis (129). Additionally, the sequential approach assumes no unmeasured mediator-mediator confounding (130). The regression models and formulae described in **3.5 Estimation of the effects under causal mediation analysis** can be used to estimate the NIEs and NDE under the sequential mediation analysis approach (the method used for analysis of data in **Chapter 7**). However, for each step in the sequence a new model for the outcome needs to be fitted. This not only increases the number of models needed to estimate the effects, it might also lead to incompatibility between outcome models, especially if mediator-mediator interaction terms are included in the models. For a binary outcome, the

regression-based approach can only be used if all the mediators are continuous (129). Additionally, even when all mediators are continuous, the expression for the NDE gets complex if the models include exposure-mediator interaction terms, because they need to account for correlation between the mediators (as shown in **3.5 Estimation of the effects under causal mediation analysis**). Either a Monte-Carlo approach (130) (similar to the approach described in section **3.7.3 Interventional analysis** or a weighting approach are able to remedy these issues (129). As described above, the regression-based approach can be modified to accommodate case-control or case-cohort data.

The weighting approach (the method used for analysis of data in **Chapter 5**) involves first estimating the inverse probability weights for the exposure from a logistic regression model for the exposure conditional on baseline confounders as below:

$$logit(P(A = a|c)) = \theta_0 + \theta_1 c$$

From this model, the inverse probability weights (IPW) for each individual i can be calculated as:

$$\frac{P(A = a)}{P(A = a|C = c_i)} \text{ for the exposed;}$$

and

$$\frac{P(A = a^*)}{P(A = a^*|C = c_i)} \text{ for the unexposed.}$$

Then, the population-level averages of the three counterfactuals needed to estimate the NDE and NIE can be estimated as follows. Quantity $E[Y_{aM_a}]$ is estimated as the weighted average of the outcome among the exposed ($A = a$) with weight equal to $\frac{P(A=a)}{P(A=a|C=c_i)}$, and $E[Y_{a^*M_{a^*}}]$ as

the weighted average of the outcome among the unexposed ($A = a^*$) with weight equal to $\frac{P(A=a^*)}{P(A=a^*|C=c_i)}$. As for, $E[Y_{aM_{a^*}}]$, first a model for the outcome has to be fitted as follows:

$$E[Y|a, m, c] = \alpha_0 + \alpha_1 a + \alpha_2 m + \alpha_3 am + \alpha_4 c$$

Then, $E[Y_{aM_{a^*}}]$ is estimated as the weighted average, with weights equal to $\frac{P(A=a^*)}{P(A=a^*|C=c_i)}$, of the predicted estimate for the outcome if an unexposed individual were exposed but their mediator and confounder values were as observed.

In the outcome model above, m can be either a single mediator or a vector of multiple mediators. In the case of multiple mediators, a similar sequential process as illustrated in Figure 3-10 can then be applied to estimate the sequential effects. While the weighting approach is able to deal with the limitations of the regression-based approach, it performs best when the exposure is binary (129). This approach does not accommodate case-control or case-cohort data.

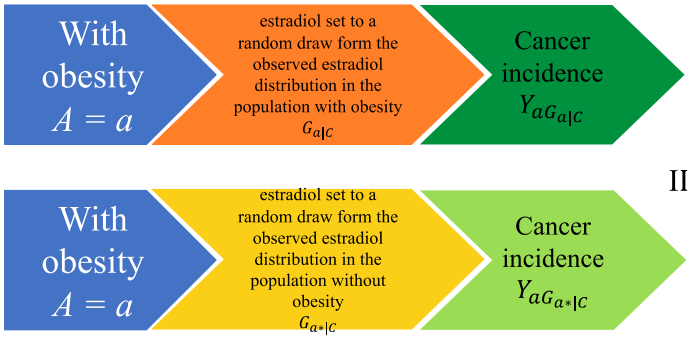
3.7.3 Interventional analysis

More recently, the causal mediation analysis literature has introduced interventional indirect (IIE) and direct (IDE) effects, which are defined based on population level interventions that could shift the observed distribution of a mediator of interest (e.g. estradiol distribution in the women with obesity given the baseline confounders) to another observed distribution (e.g. estradiol distribution in the women with normal weight given the baseline confounders) (see below for formal definitions of the IIE and IDE). Thus, contrary to the *nie* and *nde*, which are defined at the individual level and rely on unobservable counterfactual scenarios (e.g. estradiol level a woman would naturally have had she had obesity vs estradiol level same woman would naturally have she had normal weight), the IIE and IDE are defined at population level and are based on observed distributions (130-132). Importantly, the IIE and IDE can be identified if

the no confounding assumption (iv) (i.e. no exposure-induced mediator-outcome confounder, see **3.4 No confounding assumptions for mediation analysis**) is violated and therefore in the presence of multiple interrelated mediators. This is because intervening to set the mediator level to a value [randomly drawn from a distribution given the baseline confounders] essentially removes arrows from other mediators influencing the mediator level (130-132). A drawback of the IIE and IDE is that in the presence of multiple mediators interacting with each other, the TE might not completely decompose into the IIE and IDE because these effects do not capture the indirect effect through dependence between the mediators (130-132). Similar to the NIE and NDE, identifying the IIE and IDE relies on the assumptions of positivity and consistency (126). The formal definitions of IIE and IDE are as follows:

The IIE is defined as $E(Y_{aG_{a|C}} - Y_{aG_{a*|C}})$, which is the change in the outcome among the exposed, if an intervention could change the distribution of a mediator of interest from what has been observed for the exposed to what has been observed for the unexposed conditional on baseline confounders. For this effect to be non-zero, the observed distribution of the mediator needs be different in the exposed and unexposed group, and that difference must influence the outcome (Figure 3-11) (130-132).

Figure 3-11 Schematic illustration of the definition of the interventional indirect effect of adiposity, as a binary variable (with vs without obesity), on cancer incidence through estradiol

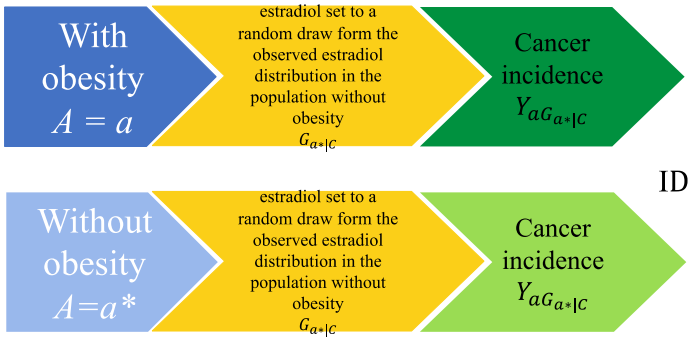


$$\text{IIE} = E[Y_{aG_{a*|C}} - Y_{a^*G_{a*|C}}]$$

Abbreviations IIE interventional indirect effect

The IDE is defined as $E(Y_{aG_{a*|C}} - Y_{a^*G_{a*|C}})$, which is the change in the outcome for the exposed compared with the unexposed, if an intervention could fix the distribution of the mediator in the exposed to the observed distribution of the mediator in the unexposed conditional on baseline confounders (130-132). This can be done by assigning a mediator value to the individuals in the exposed group, randomly drawn from the distribution of the mediator in the unexposed (Figure 3-12) (130).

Figure 3-12 Schematic illustration of the definition of the interventional direct effect of adiposity, as a binary variable (with vs without obesity), on cancer incidence not through estradiol



$$\text{IDE} = E[Y_{aG_{a*|C}} - Y_{a^*G_{a*|C}}]$$

Abbreviations IDE interventional direct effect

Two approaches, namely a "standardisation-based" and a "multiple-mediator regression-standardisation approach" approach, are available for estimating interventional effects (131). The second approach, proposed by Vansteelandt and Daniel (131) (the method used for analysis of data in **Chapter 6** and **Chapter 8**) is described below.

Under this approach, for a continuous outcome, the IIE(c) through M_1 could be defined as:

$$E\left[\sum_{m_1}\sum_{m_2}\sum_{m_3} E(Y_{am_1m_2m_3}|c)\{P(M_{1a} = m_1|c) - P(M_{1a^*} = m_1|c)\}P(M_{2a*} = m_2|c)\,P(M_{3a*} = m_3|c)\right]$$

The IIE(c) through M_2 could be defined as:

$$E\left[\sum_{m_1}\sum_{m_2}\sum_{m_3} E(Y_{am_1m_2m_3}|c)\{P(M_{2a}=m_2|c) - P(M_{2a^*}=m_2|c)\}P(M_{1a}=m_1|c)\,P(M_{3a*} = m_3|c)\right]$$

The IIE(c) through M_3 could be defined as:

$$E\left[\sum_{m_1}\sum_{m_2}\sum_{m_3} E(Y_{am_1m_2m_3}|c)\{P(M_{3a}=m_3|c) - P(M_{3a^*}=m_3|c)\}P(M_{1a}=m_1|c)\,P(M_{2a} = m_2|c)\right]$$

The IDE(c) not through M_1, M_2, M_3 could be defined as:

$$E\left[\sum_{m_1}\sum_{m_2}\sum_{m_3}\{E((Y_{am_1m_2m_3}|c) - (Y_{a^*m_1m_2m_3}|c)\}P(M_{1a*}=m_1, M_{2a*}=m_2, M_{3a*}=m_3|c)\right]$$

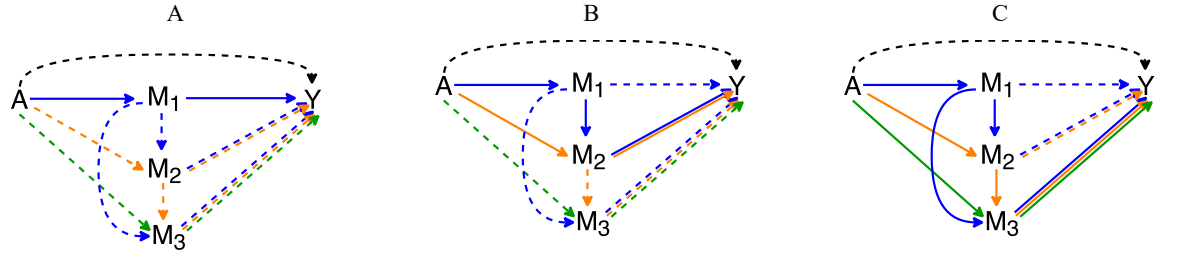
Similar effects can be defined for binary outcomes on the relative (risk ratio or odds ratio) scale (130).

If the assumed causal ordering depicted in Figure 3-13 holds, i.e. M_1 precedes M_2 and M_3 and M_2 precedes M_3, then the above decompositions correspond to the IIE through M_1 as $A \rightarrow M_1 \rightarrow Y$ (Figure 3-13**A**), the IIE through M_2 as the combination of $A \rightarrow M_1 \rightarrow M_2 \rightarrow Y$ and $A \rightarrow M_2 \rightarrow Y$ (Figure 3-13**B**), and the IIE through M_3 as the combination of $A \rightarrow M_1 \rightarrow M_2 \rightarrow M_3 \rightarrow Y$, $A \rightarrow M_2 \rightarrow M_3 \rightarrow Y$ and $A \rightarrow M_3 \rightarrow Y$ (Figure 3-13**C**). However, importantly, the IIE and IDE can be used without having to assume a causal ordering of the mediators, which compared with the sequential approach, is a considerable advantage.

Additionally, the IIE and IDE are still identifiable when the mediators are not causally ordered at all, e.g., when the dependence between them (after conditioning on *A* and *C*) is due to unobserved confounders.

To overcome the issue of the TE not completely decomposing into the IIEs and IDE, Vansteelandt and Daniel define an additional indirect effect as the difference between the TE and the IIEs and IDE. This effect would capture the indirect effect of exposure on outcome through the interdependence between mediators (130).

Figure 3-13 Diagrams illustrating the interventional indirect effects estimated using the decomposition approach proposed by Vansteelandt and Daniel (130)



In this diagram, A is the exposure, Y the outcome, and M_1 , M_2, M_3 the mediators. In diagrams A, B, and C the solid arrows represent the interventional indirect effects through M_1 , M_2, and M_3 respectively.

The IIEs and IDE can be estimated from the regression models described in **3.5 Estimation of the effects under causal mediation analysis** and averaging the estimates based on a large number of Monte-Carlo draws from the mediator distributions, with the distribution of the confounders set to their observed distribution in the sample (130).

The approaches described above for assessment of mediation in the presentation of multiple interrelation mediators have been summarised in Table 3-1.

Table 3-1 Summary of causal mediation analysis approaches for effect decomposition in the presence of multiple interrelated mediators

		Joint analysis using regression-based method	Joint analysis using weighting method	Sequential analysis using regression-based method	Sequential analysis using weighting method	Interventional analysis using regression-standardisation method
Allows for correlated Ms		✓	✓	✓	✓	✓
Agnostic to causal ordering between Ms		✓	✓			✓
Provides insight into different mediating roles of Ms				✓	✓	✓
Exposure type		Binary and continuous	Preferably binary	Binary and continuous	Preferably binary	Binary and continuous
Mediator type		Binary and continuous But for binary outcome only allows continuous Ms	Any	Binary and continuous But for binary outcome only allows continuous Ms	Any	Binary and continuous
Handles $A-M$ interaction		For binary outcome expression for NDE gets complicated	✓	For binary outcome expression for NDE gets complicated	✓	✓
Handles $M-M$ interaction		Can lead to model incompatibility issues	✓	Can lead to model incompatibility issues	✓	✓
Can be used for case-control studies		✓		✓		
Other				Assumes no unmeasured $M-M$ confounding	Assumes no unmeasured $M-M$ confounding	Computationally intensive
Models needed	$Y\mid A, Ms, C$	✓	✓	✓	✓	✓
	$A\mid C$ (to obtain propensity scores)		✓		✓	
	$M\mid A, C,$ $upstream\ Ms$	✓		✓		✓
	$M\mid A, C$					✓

A is the exposure, Y the outcome, and M is the mediator, and C is the vector for baseline confounders.
Abbreviations NDE natural direct effect

3.8 Inclusion of interaction in mediation analysis

An important advantage of the causal mediation analysis methods over the traditional approaches is that they can allow for possible exposure-mediator interaction (19). As explained in section **3.2 Traditional approaches to mediation analysis**, presupposing no exposure-mediator interaction might lead to biased estimation of the indirect and direct effects (19). Approaches for multiple mediators are also able to handle mediator-mediator interactions, and

more complex interaction terms such as exposure-mediator-mediator or mediator-mediator-mediator interactions.

There is no clear answer to the question of when and which interactions terms to include, especially as the settings get more complicated with multiple mediators. The general advice is against using a statistical test to decide about the presence of interaction, because statistical tests can be underpowered to identify an interaction, and interactions of moderate magnitude, difficult to detect in practice, can still influence the indirect and direct effects. Instead it is suggested to keep the interaction term in cases where the magnitude of the term in the outcome model is large and the estimated indirect and direct effects with and without the inclusion of the interaction term are different enough to change the conclusion from the analysis (19). This approach can be easy to follow in single mediator settings. However, as the number of mediators and as a consequence, the number of interaction terms increase, the cost of increased number of parameters in the models also needs to be considered, especially for smaller studies.

3.9 Epidemiological studies investigating mediating roles of biomarkers in the effect of adiposity on cancer risk

Colorectal Cancer

Results from prospective studies investigating the mediating effect of biomarkers on the three putative pathways linking adiposity to colorectal cancer are summarised in Table 3-2.

The difference method was used in one study for men (117) and three studies for women (59, 83, 117) to investigate the mediating effects of several biomarkers of inflammatory status (i.e. non-HMW adiponectin, sOB-R, CRP, and leptin), HbA1c, C-peptide, and fasting insulin on the association between waist circumference and colon or colorectal cancer (Table 3-2). Findings from these studies suggest that inflammatory status biomarkers, except CRP, have

some mediating role in the association. Also, in women, C-peptide or fasting insulin explained some (~30%) of the association (Table 3-2).

A causal mediation analysis approach was used to quantify the mediating effect of a panel of inflammatory and metabolic biomarkers jointly (CRP, IL-6, TNF-receptor 2 (TNFR-2), macrophage inhibitory cytokine-1 (MIC-1), adiponectin, sOB-R, and C-peptide) in a case-control study nested within the HPFS (using a joint mediation analysis approach similar to the one described in **3.7.1 Joint analysis**). The OR per 3.6 kg/m^2 increment in BMI for the total BMI-colorectal cancer association and the NIE through all the biomarkers were 1.40 (95%CI, 1.14-1.73) and 1.26 (95%CI, 0.97 – 1.52), respectively (133). The observed NIE through CRP, IL-6, TNFR-2, and MIC-1 combined was smaller than the NIE effect through adiponectin, sOB-R, and C-peptide (Table 3-2). This latter analysis allowed for possible correlation between biomarker within each set but assumed the two sets of biomarkers were independent of each other. In analyses of individual biomarkers, assuming independence between biomarkers, there was no evidence for a mediating effect through any single biomarker, but the estimated NIEs were larger for adiponectin and sOB-R compared with other assessed biomarkers (133). The estimated indirect effects were not replicated when waist circumference was used as the measure of adiposity (Table 3-2) (133).

In another recent causal mediation analysis, triglyceride-glucose index (TyG; a proxy for insulin resistance) weakly mediated the associations between BMI and colon (PM 20%) and rectal cancer (PM 34%) (Table 3-2). No other biomarkers were assessed as potential mediators in this study (134).

Table 3-2 Total, direct, and indirect effects for the association between adiposity and colorectal, colon, or rectal cancers estimated from cohort, nested case-control or case cohort studies

Study (Ref)	Biomarker	N participants	Population	Adiposity measure/ Comparison	[measure] Total effect (95% CI)	Direct effect (95%CI) Or Adiposity-cancer association from models adjusted for the mediator(s) (95%CI)	Indirect effect (95%CI) Or % attenuated
Difference method *							
Difference method Matched nested case-control EPIC (117)	Non-HMW Adiponectin	292 colon cancer cases, 292 controls	Men	WC / 10 cm	[OR] 1.33 (1.26 – 1.41)	1.27 (1.20 – 1.35)	16%
	sOB-R					1.21 (1.14 – 1.29)	33%
	CRP-hs					1.31 (1.24 – 1.39)	6%
	HbA1c					1.31 (1.24 – 1.39)	6%
	C-peptide					1.37 (1.29 – 1.45)	-9%
	Non-HMW Adiponectin	370 colon cancer cases, 370 controls	Women	WC / 10 cm	[OR] 1.20 (1.14 – 1.25)	1.14 (1.08 – 1.19)	28%
	sOB-R					1.14 (1.09 – 1.19)	28%
	CRP-hs					1.19 (1.14 – 1.25)	2%
	HbA1c					1.18 (1.13 – 1.24)	4%
	C-peptide					1.13 (1.12 – 1.23)	30%
Case-cohort WHI-OS (59)	Leptin	457 colorectal cancer cases, 841 sub-cohort	Postmenopausal women	WC / 1 cm	[HR] 1.02 (1.01 – 1.04)	Not reported	25%
	Fasting insulin					Not reported	37%
	Leptin, fasting insulin**					1.01 (1.00 – 1.03)	50%
Case-cohort WHI-OS (83)	Fasting insulin	438 colorectal cancer cases, 809 sub-cohort	Postmenopausal women without diabetes	WC/ Q4 vs Q1	[HR] 1.82 (1.22 – 2.70)	1.48 (0.94 – 2.32)	35%
Causal mediation analysis							
Matched nested case-control HPFS (133)	Adiponectin***	209 colorectal cancer cases, 382 controls	Men	BMI/ 3.6 kg/m^2	[OR] 1.40 (1.13 – 1.73)	1.28 (1.04 – 1.59)	1.09 (0.93 – 1.25)
	sOB-R***					1.25 (0.99 – 1.58)	1.12 (0.95 – 1.29)
	CRP***					1.35 (1.07 – 1.69)	1.04 (0.96 – 1.13)
	IL-6***					1.36 (1.09 – 1.69)	1.03 (0.99 – 1.08)
	TNFR-2***					1.41 (1.14 – 1.75)	0.99 (0.98 – 1.01)
	MIC-1***					1.40 (1.13 – 1.74)	1.00 (0.99 – 1.01)
	C-peptide***					1.33 (1.06 – 1.68)	1.05 (0.91 – 1.18)
	All above****					1.11 (0.87 – 1.42)	1.26 (0.97 – 1.52)
	CRP, IL6, TNFR-2, MIC****					1.34 (1.07 – 1.68)	1.05 (0.96 – 1.14)
	Adiponectin, sOB-R, C-peptide****					1.13 (0.90 – 1.41)	1.24 (0.92 – 1.55)

continued

Table 3-2 continued Total, direct, and indirect effects for the association between adiposity and colorectal, colon, or rectal cancer estimated from cohort, nested case-control or case cohort studies

Study (Ref)	Biomarker	N participants	Population	Adiposity measure/ Comparison	Total effect (95% CI) measure	Direct effect (95%CI) Or Adiposity-cancer association from models adjusted for the mediator(s) (95%CI)	Indirect effect (95%CI) Or % attenuated
				Causal mediation analysis			
Matched nested case-control HPFS (133)	Adiponectin***	166 colorectal cancer cases, 318 controls	Men	WC/ 11.4 cm	[OR] 1.71 (1.30 – 2.25)	1.74 (1.28 – 2.37)	0.98 (0.82 – 1.15)
	sOB-R***					1.79 (1.32 – 2.43)	0.96 (0.82 – 1.12)
	CRP***					1.75 (1.31 – 2.35)	0.98 (0.90 – 1.06)
	IL-6***					1.70 (1.28 – 2.27)	1.00 (0.95 – 1.06)
	TNFR-2***					1.76 (1.33 – 2.33)	0.97 (0.93 – 1.01)
	MIC-1***					1.70 (1.29 – 2.24)	1.01 (0.98 – 1.04)
	C-peptide***					1.80 (1.31 – 2.46)	0.95 (0.80 – 1.13)
	All above****					1.90 (1.35 – 2.66)	0.90 (0.70 – 1.16)
	CRP, IL6, TNFR-2, MIC****					1.77 (1.31 – 2.38)	0.97 (0.89 – 1.05)
	Adiponectin, sOB-R, C-peptide****					1.84 (1.32 – 2.57)	0.93 (0.72 – 1.18)
Six European cohorts (134)	TyG index***	510,471 (4032 affected with colon cancer)	Men and women	BMI/ 5 kg/m^2	[RR] 1.14 (1.10 – 1.19)	1.11 (1.07 – 1.16)	1.03 (1.01 – 1.04)
		510,471 (2430 affected with rectal cancer)			[RR] 1.09 (1.03 – 1.15)	1.06 (1.00 – 1.12)	1.03 (1.01 – 1.05)

Abbreviations HMW high molecular weight; sOB-R soluble leptin receptor; CRP-hs high sensitivity C-reactive protein; HbA1c glycosylated haemoglobin; IL interleukin; TNFR tumour necrosis factor receptor; MIC macrophage inhibitory cytokine-1; TyG triglyceride glucose product; WC waist circumference; BMI body mass index; CI confidence interval; OR odds ratio; HR hazard ratio; RR risk ratio; Ref reference; EPIC European Prospective Investigation into Cancer and Nutrition; WHI-OS Women's Health Initiative Observational Study; HPFS Health Professional Follow up Study

*In analyses using the difference method, each biomarker (or group of biomarkers, for example denoted by ** in the table) were included in the models for adiposity-cancer association one at a time.

*** These causal mediation analyses quantified the mediating effect of each mediator one at a time, assuming independence between the biomarkers.

**** These causal mediation analyses quantified the mediating effect through the listed biomarkers jointly, using an approach similar to the one described in section ***3.7.1***

Postmenopausal breast Cancer

Results from prospective studies investigating the mediating effect of biomarkers on pathways linking adiposity to postmenopausal breast cancer are summarised in Table 3-3.

In four studies, the difference method was used to explore the mediating effects of sex-steroid hormones (estradiol, estrone, testosterone), SHBG, CRP, and fasting insulin on the effect of adiposity on postmenopausal breast cancer (16, 68, 119, 121). Overall, the studies found some mediating effect through sex-steroid hormones, with the largest attenuation in the association observed after adjusting for free estradiol (16, 121). Also, in the WHI-OS study with available measures for CRP and fasting insulin, adjusting for each biomarker explained about 50% of the BMI-postmenopausal breast cancer association (Table 3-3) (68, 119).

In a case cohort study within the WHI-OS, the indirect effect through estradiol and insulin was quantified using a causal mediation analysis approach assuming the biomarkers were independent. The analysis showed that of the total BMI-postmenopausal breast cancer association, 24% was mediated through estradiol, with a larger mediating effect (49%) observed for ER-positive breast cancers (120). In women without diabetes, 66% of the total effect was mediated through fasting insulin (Table 3-3) (120).

In a case-control study nested within the PLCO, the estimated NIE effect of BMI on ER-positive postmenopausal breast cancer risk through unconjugated estradiol corresponded to 20% PM of the total effect (135). No other pathway was assessed in this analysis.

In another causal mediation analysis, TyG (a proxy for insulin resistance) had a negligible mediating effect on the association between BMI and postmenopausal breast cancer (16% PM) (134). This was the only biomarker assessed in this analysis.

Table 3-3 Total, direct, and indirect effects for the association between adiposity and postmenopausal breast cancer estimated from cohort, nested case-control or case cohort studies

Study (Ref)	Biomarker	N participants	Population	Adiposity measure/ Comparison	[measure] Total effect (95% CI)	Direct effect (95%CI) Or Adiposity-cancer association from models adjusted for the mediator(s) (95%CI)	Indirect effect (95%CI) Or % attenuated
				Difference method*			
Case-cohort WHI-OS (68)	CRP	875 cases, 821 sub-cohort	Postmenopausal women not using HT at blood draw	BMI/ ≥30.0 vs <25 kg/m^2	[HR] 1.87 (1.25 – 2.80)	1.34 (0.82 – 2.18)	53%
	Fasting insulin					1.35 (0.85 – 2.17)	52%
	Estradiol					1.69 (1.11 – 2.59)	16%
	CRP and fasting insulin**					1.05 (0.61 – 1.81)	92%
	CRP, fasting insulin and estradiol**					0.91 (0.52 – 1.61)	115%
Case-cohort WHI-OS (119)	Insulin	841 cases, 810 sub-cohort	Postmenopausal women without diabetes not using HT at blood draw	BMI/ ≥30.0 vs 18.05-<25.0 kg/m^2	[HR] 2.12 (1.26 – 3.58)	1.50 (0.80 – 2.83)	46%
	Estradiol					1.91 (1.11 – 3.27)	14%
	Insulin, estradiol, IGF-1**					1.36 (0.70 – 2.67)	59%
Nine cohort studies (16)	Estradiol	606 cases, 1,440 controls	Postmenopausal women	BMI/ 5 kg/m^2	[RR] 1.18 (1.06 – 1.31)	1.07 (0.95 – 1.20)	59%
	Free estradiol	443 cases, 906 controls			1.19 (1.05 – 1.34)	1.02 (0.89 – 1.17)	89%
	Estrone	425 cases, 932 controls			1.20 (1.06 – 1.35)	1.10 (0.96 – 1.25)	48%
	Testosterone	569 cases, 1,371 controls			1.17 (1.05 – 1.31)	1.13 (1.02 – 1.27)	22%
	SHBG	329 cases, 906 controls			1.13 (0.98 – 1.32)	1.09 (0.93 – 1.28)	29%
Matched nested case-control EPIC (121)	Estradiol	613 cases, 1,139 controls	Postmenopausal women not using HT at blood draw	BMI/ 5 kg/m^2	[OR] 1.11 (0.99 – 1.25)	1.03 (0.91 – 1.16)	72%
	Free estradiol					0.99 (0.87 – 1.12)	110%
	Estrone					1.04 (0.92 – 1.17)	62%
	Testosterone					1.10 (0.98 – 1.24)	9%
	SHBG					1.07 (0.94 – 1.21)	35%

continued

Table 3-3 continued Total, direct, and indirect effects for the association between adiposity and postmenopausal breast cancer estimated from cohort, nested case-control or case cohort studies

Study (Ref)	Biomarker	N participants	Population	Adiposity measure/ Comparison	Total effect (95% CI)	Direct effect (95%CI) Or Adiposity-cancer association from models adjusted for the mediator(s) (95%CI)	Indirect effect (95%CI) Or % attenuated
Difference method*							
Matched nested case-control EPIC (121)	Estradiol Free estradiol Estrone All estrogens Testosterone All androgens SHBG	613 cases, 1,139 controls	Postmenopausal women not using HT at blood draw	WC/ 10 cm	[OR] 1.12 (1.02 – 1.24)	1.06 (0.96 – 1.17) 1.02 (0.92 – 1.14) 1.07 (0.96 – 1.18) 1.03 (0.92 – 1.15) 1.12 (1.01 – 1.23) 1.09 (0.98 – 1.22) 1.08 (0.97 – 1.21)	49% 82% 40% 74% 0% 24% 32%
Causal mediation***							
Case-cohort WHI-OS (120)****	Fasting insulin	382 cases, 409 sub-cohort	Postmenopausal women without diabetes not using HT at blood draw	BMI/ 5 kg/m^2	[extra cases per 100,000 per year] 52.0 (12.1 – 91.3)	6.9 (-46.6 – 61.1)	34.2 (9.4 – 59.0)
	Estradiol	601 cases, 1,000 sub-cohort	Postmenopausal women not using HT at blood draw	BMI/ 5 kg/m^2	50.0 (23.2 – 76.6)	38.0 (8.1 – 68.1)	11.9 (1.7 – 22.6)
		401 ER-positive cases, 1,000 sub-cohort			33.0 (10.3 -55.5)	16.9 (-7.4 – 4.11)	16.1 (7.5 – 25.2)
Six European cohorts (134)	TyG index	510,471 (3,472 affected with postmenopausal breast cancer)	women	BMI/ 5 kg/m^2	[RR] 1.05 (1.01 – 1.09)	1.04 (1.00 – 1.09)	1.01 (0.99 – 1.02)

continued

Table 3-3 continued Total, direct, and indirect effects for the association between adiposity and postmenopausal breast cancer estimated from cohort, nested case-control or case cohort studies

Study (Ref)	Biomarker	N participants	Population	Adiposity measure/ Comparison	Total effect (95% CI)	Direct effect (95%CI) Or Adiposity-cancer association from models adjusted for the mediator(s) (95%CI)	Indirect effect (95%CI) Or % attenuated
Causal mediation* **							
Matched nested case-control PLCO (135)	Unconjugated estradiol	268 ER-positive cases, 193 controls	Postmenopausal women never used HT	BMI/ 5 kg/m^2	[extra cases per 100,000 per year] 7.0 (2.1 – 11.9)	5.6 (0.1 – 11.0)	1.4 (-1.0 – 3.9)

Abbreviations CRP C-reactive protein; IGF insulin-like growth factor; SHBG sex hormone binding globulin; TyG triglyceride glucose product; ER estrogen receptor; HT hormone therapy; WC waist circumference; BMI body mass index; CI confidence interval; HR hazard ratio; RR risk ratio; OR odds ratio; Ref reference; EPIC European Prospective Investigation into Cancer and Nutrition; WHI-OS Women's Health Initiative Observational Study; PLCO Prostate, Lung, Colorectal, and Ovarian Cancer Screening Trial

*In analyses using the difference method, each biomarker (or group of biomarkers, for example denoted by ** in the table) were included in the models for adiposity-cancer association one at a time.

All causal mediation analyses were single mediation analysis. In the only study where the mediating effect of more than one biomarker was assessed (estradiol and fasting insulin, study denoted by *), the analysis assumed independence between the biomarkers.

Endometrial Cancer

Results from prospective studies investigating the mediating effect of biomarkers on pathways linking adiposity to endometrial cancer are summarised in Table 3-4.

Within the WHI-OS and EPIC, four studies investigated the mediating effects of biomarkers on the inflammatory pathway (CRP, IL-6, IL-1Ra), estrogens (estradiol and estrone), and hyperinsulinemia (fasting insulin, C-peptide) on BMI-endometrial cancer association (76, 77, 94, 114) using the difference method. Results from these studies suggested that biomarkers on each of the pathways had some mediating role on the association. Also, simultaneous adjustment for biomarkers from the three pathways (e.g. CRP, fasting insulin, and estradiol or IL-1Ra, C-peptide, and estrone) explained a considerable proportion (~70% up to 117%) of the association (Table 3-4).

In a nested case-control study within the EPIC, a principal component factor analysis approach was used to define the following component factors: "inflammation factor" with >|0.40| loading for TNF receptor 1 and 2, IL-6, and CRP; "insulin resistance/ metabolic syndrome (IR/MS) factor" with >|0.40| loading for CRP, high density lipoprotein cholesterol, C-peptide, adiponectin, insulin-like growth factor binding protein 1 and 2, and SHBG; and "steroid factor" with >|0.40| loading for testosterone, DHEAS, androstenedione, estrone, and estradiol. Using the difference method to explore the mediating effects of each of these factors on BMI-endometrial cancer association, it was observed that adjusting for the IR/MS factor attenuated the association by 50%, while inflammation and steroid factors had minimal effects on the association (118) (Table 3-4).

Using causal mediation analysis, the estimated NIE of BMI on endometrial cancer risk through TyG was essentially null (~2% PM) suggesting that biomarker did not explain any of the total BMI-endometrial cancer risk association (134).

Table 3-4 Total, direct, and indirect effects for the association between adiposity and endometrial cancer estimated from cohort, nested case-control or case cohort studies

Study (Ref)	Biomarker	N participants	Population	Adiposity measure/ Comparison	[measure] Total effect (95% CI)	Direct effect (95%CI) Or Adiposity-cancer association from models adjusted for the mediator(s) (95%CI)	Indirect effect (95%CI) Or % attenuated
				Difference method *			
Case-cohort WHI-OS (76)	CRP Fasting insulin Estradiol CRP, fasting insulin and estradiol**	151 cases, 301 sub-cohort	Postmenopausal Women not using HT at blood draw	BMI / ≥30.0 vs <25 kg/m^2	[HR] 1.85 (1.13 – 3.04)	1.23 (0.66 – 2.29) 1.15 (0.63 – 2.11) 1.38 (0.80 – 2.39) 0.90 (0.46 – 1.78)	67% 77% 48% 117%
Matched nested case-control EPIC (77)	IL-6 IL-1Ra CRP C-peptide Estrone IL-6, C-peptide, estrone** IL-1Ra, C-peptide, estrone** CRP, C-peptide, estrone**	305 cases, 574 controls	Pre and postmenopausal women	BMI / ≥30.0 vs <25 kg/m^2	[OR] 2.02 (1.26 – 3.23)	1.72 (1.04 – 2.85) 1.65 (1.01 – 2.71) 1.79 (1.06 – 3.02) 1.43 (0.86 – 2.40) 1.75 (1.07 – 2.85) 1.16 (0.67 – 2.01) 1.10 (0.64 – 1.90) 1.21 (0.69 – 2.14)	23% 29% 17% 49% 20% 79% 86% 73%
Matched nested case-control EPIC (94)	C-peptide	286 cases, 555 controls	Pre and postmenopausal women not using exogenous hormones at blood draw	BMI / ≥30.0 vs <25 kg/m^2	[OR] 2.37 (1.60 – 3.51)	1.93 (1.27 – 2.92)	24%
Matched nested case-control WHI-OS (114)	Unconjugated estradiol	313 cases, 354 controls	Postmenopausal Women not using HT at blood draw	BMI / ≥30.0 vs <25 kg/m^2	[OR] 3.97 (2.54 – 6.21)	2.25 (1.33 – 3.81)	41%

continued

Table 3-4 continued Total, direct, and indirect effects for the association between adiposity and endometrial breast cancer estimated from cohort, nested case-control or case cohort studies

Study (Ref)	Biomarker	N participants	Population	Adiposity measure/ Comparison	[measure] Total effect (95% CI)	Direct effect (95%CI) Or Adiposity-cancer association from models adjusted for the mediator(s) (95%CI)	Indirect effect (95%CI) Or % attenuated
				Difference method*			
Matched nested case-control EPIC (118)	Inflammation factor*†	183 cases, 452 controls	Postmenopausal women not using HT at blood draw	BMI / ≥30.0 vs <25 kg/m^2	[OR] 2.73 (1.66 – 4.50)	2.55 (1.51 – 4.29)	7%
	Insulin resistance/ metabolic syndrome factor*†					1.65 (0.92 – 2.98)	50%
	Steroids factor*†					2.72 (1.65 – 4.50)	4%
				Causal mediation***			
Six European cohorts (134)	TyG index	510,471 (1,417 affected with endometrial cancer)	Women	BMI/ 5 kg/m^2	[RR] 1.50 (1.43 – 1.57)	1.49 (1.41 – 1.57)	1.01 (0.99 – 1.03)

Abbreviations CRP C-reactive protein; IL interleukin; IL-1Ra interleukin 1 receptor antagonist; TyG triglyceride glucose product; HT hormone therapy; BMI body mass index; CI confidence interval; HR hazard ratio; OR odds ratio; RR risk ratio; Ref reference; WHI-OS Women's Health Initiative Observational Study; EPIC European Prospective Investigation into Cancer and Nutrition

*In analyses using the difference method, each biomarker (or group of biomarkers, for example denoted by ** in the table) were included in the models for adiposity-cancer association one at a time.

*** The causal mediation analyses were single mediation analysis.

†Factors were defined using principle components factor analysis. For factor named "inflammation", tumour necrosis factor receptor 1 and 2, interleukin 6, and C-reactive protein had loading >|0.40|. For factor named "Insulin resistance/ metabolic syndrome factor", C-reactive protein, high density lipoprotein cholesterol, C-peptide, adiponectin, insulin-like growth factor binding protein 1 and 2, and sex hormone binding globulin had loading >|0.40|. For factor named "steroids" testosterone, DHEAS, androstenedione, estrone, and estradiol had loading >|0.40|.

3.10 Summary

Mediation analysis methods allow assessing the contribution of mediating factors to the effect of exposure on outcome. In this chapter, an overview of these approaches was provided. First, the traditional approaches for mediation analysis, namely the difference and the product methods, and their limitations were described. Although intuitive and straightforward, both of these methods presuppose no exposure-mediator interaction, and linear exposure-outcome and mediator-outcome associations. Also, they do not extend to situations where there are multiple interrelated mediators connecting the exposure to the outcome. The chapter then described the causal mediation analysis methods. These methods allow quantification of the mediated effects in the presence of exposure-mediator interactions and non-linearities and thus supersede the traditional approaches. More recently extensions of these have become available that allow effect decomposition in the presence of multiple interrelated mediators. These approaches, which have been applied in subsequent chapters, were described and summarised in Table 3-1.

In the final section, epidemiological studies investigating the mediating effects of biomarkers reflecting the three pathways of interest (i.e. inflammatory status, insulin signalling, and sex-steroid hormones) in the effect of adiposity on colorectal, postmenopausal breast, and endometrial cancers were reviewed. Most of the studies used the difference method to investigate the change in the measure of association between adiposity and cancer risk after adjusting for the biomarker of interest and did not quantify the indirect effect through the biomarker. Four causal mediation analysis studies were found. In two, only a single mediator was assessed. In one, both insulin and estradiol were evaluated, but the analysis assumed the mediators were independent. In another, a panel of inflammatory markers and C-peptide were assessed jointly.

The identified studies offer some clues about the potential mediating role of biomarkers in the effect of adiposity on cancer risk. For colorectal cancer, inflammatory biomarkers, insulin, and C-peptide seem to explain some of the effect, but studies assessing the mediating role of sex-steroid hormones are lacking. For postmenopausal breast cancer, there is evidence that estradiol, especially free estradiol, plays a major mediating role, but one study suggests even a more important mediating role for fasting insulin. For endometrial cancer, the difference method suggests that estradiol might explain the effect only in part and other biomarkers, including inflammatory biomarkers, fasting insulin or C-peptide might also be involved.

Studies attempting to quantify the mediating role of the three pathways postulated to connect adiposity to cancer risk simultaneously are limited and even fewer have endeavoured to disentangle the pathways to assess their possibly different mediating roles.

Chapter 4 Methods

This Chapter is a brief overview of the four cohort studies used in this thesis and the statistical mediation analysis methods employed for each. Detailed descriptions are provided in Chapters 5 to 8.

4.1 Studies

This thesis made use of data from four existing studies.

i) The UK Biobank cohort (136) to investigate the mediating roles of C-reactive protein (CRP, to reflect the inflammatory status), glycosylated haemoglobin (HbA1c, to reflect insulin resistance), and sex hormone-binding globulin (SHBG) and testosterone (to reflect alterations in sex-steroid hormone levels) in the effect of adiposity on colorectal cancer risk in men and postmenopausal women (**Chapter 5**)

ii) A case-cohort study within the Melbourne Collaborative Cohort Study (MCCS) (137, 138) to investigate the mediating roles of fasting insulin (to reflect hyperinsulinemia) and calculated free estradiol (to reflect alterations in sex-steroid hormone levels) in the effect of adiposity on estrogen receptor (ER)-positive breast cancer risk in postmenopausal women (**Chapter 6**)

iii) A case-control study nested within the European Prospective Investigation into Cancer and Nutrition (EPIC) (77, 78, 94, 113, 139, 140) to investigate the mediating roles of adiponectin, interleukin-6 (IL-6), interleukin-1-receptor antagonist (IL-1Ra), TNF-receptor-1 (TNF-R1), TNF-R2, and CRP (to reflect the inflammatory status), C-peptide (to reflect hyperinsulinemia), and calculated free estradiol and estrone (to reflect alterations in sex-steroid hormone levels) in the effect of adiposity on endometrial cancer risk in postmenopausal women (**Chapter 7**).

iv) A case-cohort study within the Women's Health Initiative Observational Study (WHI-OS) (59, 68, 76, 83, 90, 119, 141) to investigate the mediating roles of adiponectin, leptin, resistin, hepatocyte growth factor (HGF), plasminogen activator inhibitor (PAI), IL-6, and CRP (to reflect the inflammatory status), fasting insulin (to reflect hyperinsulinemia), and total estradiol (to reflect alterations in sex-steroid hormone levels) in the effect of adiposity on ER-positive breast, endometrial, and colorectal cancer risk in postmenopausal women (**Chapter 8**).

An overview of these cohort studies is shown in Table 4-1**,** and a detailed description of each appears in the relevant chapters.

Table 4-1 Overview of the characteristics of the four cohort studies included in this thesis

Cohort/ related chapter in thesis	Rcrt years	Rcrt location	Demographic characteristics	N participants	Follow-up and cancer ascertainment	Collection of baseline information Anthropometric Measures (waist, hip, height, and weight)	Blood sample for biomarker measurement	Demographics, lifestyle, health history	Cohort profile ref
UK Biobank / Chapter 5	2006 to 2010	England, Wales, and Scotland	54% women 40 – 69 yo	502,506*	Linkage to the UK National Health Service	Directly measured for almost all participants	Plasma and serum stored at -80°C or in liquid nitrogen (fasting not required)	Self-administered touchscreen questionnaire	(136)
MCCS / Chapter 6	1990 to 1994	Melbourne metropolitan area	59% women 40 – 69 yo	41,513	Linkage to Victorian death records, Victorian Cancer Registry and the National Death Index	Directly measured for almost all participants	Plasma stored in liquid nitrogen (74% fasting)	Interviewer administered questionnaire	(137)
EPIC / Chapter 7	1992 to 2000	23 centres from France, Germany, Greece, Italy, The Netherlands, Spain, The UK, Sweden Denmark Norway	70% women 35 – 70 yo	519,978	*France, Germany, Greece*: Combination of active follow-up, health insurance records, cancer and pathology registries *Other countries*: linkage to population cancer registries	*France and part of UK*: all self-reported *All other centres*: directly measured *One centre in Sweden, one in Norway*: only measured weight and height	Plasma and serum *Sweden:* stored at -80°C *Denmark:* stored in nitrogen vapour (-150°C) *All other centres:* stored in liquid nitrogen (-196°C) (fasting not required)	Self or interviewer administered questionnaire	(139)
WHI-OS / Chapter 8	1993 to 1998	40 centres across the US	Post-menopausal women 50 – 79 yo	93,676	Active follow-up via annual self-administered questionnaires, confirmed using centralized review of medical reports	Directly measured for almost all participants	Plasma and serum stored at -70°C (all fasting)	Self-administered questionnaire	(141)

Abbreviations MCCS Melbourne Collaborative Cohort Study; EPIC European Prospective Investigation into Cancer and Nutrition; WHI-OS Women's Health Initiative Observational Study; Rcrt Recruitment; yo years old; ref reference

*the number is based on the reported number retrieved from the website (http://biobank.ndph.ox.ac.uk/showcase/field.cgi?id=31) on 11th February. This number is updated frequently as some participants withdraw consent.

The end of follow-up of the studies, the eligibility criteria imposed on them for the analyses presented in this thesis, and the previously published work related to the sub-studies are provided in Table 4-2.

For all analyses, participants classified as underweight (BMI<18.5) and those with a history of cancer diagnosis at recruitment (baseline) were excluded. Participants who at baseline had a history of diagnosis with diabetes (and diabetes medication use, when this information was available, i.e. in the UK Biobank and the MCCS) were also excluded because these might influence the association between adiposity and insulin signalling. For women, all analyses were limited to postmenopausal women because the mechanistic pathways, especially the sex-steroid hormone pathway, might be different before and after the menopause. Also, women who at baseline reported using hormone therapy were not included in the analyses because of the possible influence of exogenous hormones on measured sex-steroid hormone levels. For analyses with endometrial cancer as the outcome, women who at baseline were not at risk of that cancer due to a history of hysterectomy were also excluded. To reduce the possibility of [preclinical] cancer having influenced adiposity or levels of biomarkers, cancer cases diagnosed within the first year after baseline were not included in the analysis of the UK Biobank, MCCS, and WHI-OS data. Similarly, participants unaffected with cancer who had less than one year of follow up were excluded. For the EPIC study, all cases were included in the primary analysis due to the small study size but the results of the analysis excluding cases diagnosed within the first two years after baseline and controls with less than two years of follow up were presented as a sensitivity analysis.

Despite the designs of the UK Biobank study (cohort), and MCCS and WHI-OS sub-studies (case-cohort), cancer diagnosis was not treated as a time to event outcome in the mediation analyses for these studies (see **4.2 Statistical analysis** for additional explanation). Therefore, the analytic dataset for mediation analyses excluded participants who died or were lost to follow-up before the end of the follow-up period.

A list of the measured biomarkers within the four studies that were considered as potential mediators in this thesis, the assays used to measure them, their unit of measurement, and the total, between, or within assay coefficients of variation (CV), as reported in the previous publications based on the same data, are shown in Table 4-3. For the UK Biobank and MCCS, where a subset of participants had measurements available for two time points, the intraclass correlation coefficients (ICCs) are also reported. A brief description of the methods used for biomarker measurement in each study also appears in the relevant chapters.

Table 4-2 Designs of the sub-studies within each cohort this thesis made use of, the cancer outcome the studies contributed and their end of follow-up, and the eligibility criteria imposed on each study by the analyses presented in this thesis

Cohort/ related chapter in thesis	Design of the sub-study	Outcome	End of follow-up	Eligibility criteria	Ref related to the sub-study
UK Biobank / Chapter 5	NA (full cohort data was used)	Colorectal cancer in men and postmenopausal women	March 2016 for England and Wales October 2015 for Scotland	At baseline: -Body size measurements collected -BMI≥18.5 kg/m^2 -Not diagnosed with cancer -No history of diabetes or diabetes medication use - >1-year follow-up (cancer cases diagnosed >1 year after baseline) -No missing values for potential baseline confounders Additionally, for cancer unaffected: -Unaffected participants not died or lost to follow-up before the end of follow-up Additionally, for cancer affected: -Not diagnosed with appendix cancer Additionally, for women: -Postmenopausal* -Not using hormonal contraceptive or hormone therapy	NA
MCCS / Chapter 6	Case-cohort	ER-positive breast cancer in postmenopausal women	June 2002	At baseline: -Body size measurements and blood samples collected -BMI≥18.5 kg/m^2 -Not diagnosed with cancer -No history of diabetes or diabetes medication use ->1-year follow-up (cancer cases diagnosed >1 year after baseline) -No missing values for potential baseline -Postmenopausal** -Not using hormone therapy confounders Additionally, for mediation analysis: Cancer unaffected women who did not die and were not lost to follow-up before the end of follow-up	(138)

continued

Table 4-2 continued Designs of the sub-studies within each cohort this thesis made use of, the cancer outcome the studies contributed and their end of follow-up, and the eligibility criteria imposed on each study by the analyses presented in this thesis

Cohort/ related chapter in thesis	Design of the sub-study	Outcome	End of follow-up	Eligibility criteria	Ref related to the sub-study
EPIC / Chapter 7	Nested, matched case-control	Endometrial cancer in postmenopausal women	Between December 1999 and November 2003	At baseline: -BMI≥18.5 kg/m^2 -Not diagnosed with cancer -No history of diabetes -No missing values for potential baseline confounders -Postmenopausal*** -Not using oral contraceptives or hormone therapy -No history of hysterectomy	(77, 78, 94, 113, 140)
WHI-OS / Chapter 8	Case-cohort	ER-positive breast, endometrial, and colorectal cancer in postmenopausal women****	February 2004	At baseline: -BMI≥18.5 kg/m^2 -Not diagnosed with cancer -No history of diabetes -Not using hormone therapy -No missing values for body size measurements, potential baseline confounders, or biomarkers - >1-year follow-up (cancer cases diagnosed >1 year after baseline) Additionally, for endometrial cancer: No history of hysterectomy Additionally, for mediation analysis: Unaffected women who did not die and were not lost to follow-up before the end of follow-up	(59, 68, 76, 83, 90, 119)

Abbreviations: MCCS Melbourne Collaborative Cohort Study; EPIC European Prospective Investigation into Cancer and Nutrition; WHI-OS Women's Health Initiative Observational Study; NA not applicable; ER estrogen receptor; BMI body mass index

*Menopause status defined as self-reported postmenopausal or age >55 years old with unknown menopause status

**Menopause status defined as self-reported postmenopausal or age ≥56 years old with unknown menopause status or age <56 years old but measured total estradiol concentration <109 pmol/L

***Menopause status defined as self-reported postmenopausal or having had bilateral ovariectomy, or age >55 years old with unknown menopause status

****The study only recruited healthy postmenopausal women aged 50 – 79 years old

Table 4-3 Available biomarker measurements in each study that were considered as potential mediators of the effect of adiposity on cancer risk

Cohort/ related chapter in thesis	Biomarker (ref) *(pathway)*	Measured in	Analysis method and platform, and manufacture	Unit	Coefficient of variation (CV)/ Intraclass correlation coefficient (ICC)
UK Biobank/ Chapter 5	High sensitivity C-reactive protein (142) *(inflammatory status pathway)*	Serum	Immuno-turbidimetric; Beckman Coulter AU5800, Beckman Coulter, Ltd B, UK	mg/L	Total CVs at: IQC level low: 2.31% IQC level medium: 1.70% IQC level high: 1.69% ICC* for men 0.56 (95%CI, 0.55 – 0.57) ICC* for women 0.67 (95%CI, 0.66 – 0.68)
	Glycosylated haemoglobin (143) *(insulin signalling pathway)*	Red Blood Cells (plasma depleted EDTA samples)	High-performance liquid chromatography, Bio-Rad Variant II Turbo; Bio-Rad Laboratories, Hercules, USA	mmol/mol	Total CVs at: IQC level low: 2.13% IQC level medium: - IQC level high: 1.46% ICC* for men 0.84 (95%CI, 0.83 – 0.84) ICC* for women 0.79 (95%CI, 0.79 – 0.80)
	Sex hormone binding globulin (142) *(sex steroid hormone pathway)*	Serum	Chemiluminescent immunoassay 2-step sandwich, Beckman Coulter DXI 800; Beckman Coulter, Ltd B, UK	nmol/L	Total CVs at: IQC level low: 5.67% IQC level medium: 5.25% IQC level high: 5.22% ICC* for men 0.86 (95%CI, 0.86 – 0.87) ICC* for women 0.81 (95%CI, 0.80 – 0.82)
	Testosterone (142) *(sex steroid hormone pathway)*	Serum	Chemiluminescent immunoassay competitive binding, Beckman Coulter DXI 800; Beckman Coulter, Ltd B, UK	nmol/L	Total CVs at: IQC level low: 8.35% IQC level medium: 3.66% IQC level high: 4.15% ICC* for men 0.59 (95%CI, 0.58 – 0.61) ICC* for women 0.63 (95%CI, 0.62 – 0.64)
MCCS / Chapter 6	Fasting insulin *(insulin signalling pathway)*	Plasma	Microparticle Enzyme Immunoassay, AxSYM, Abbott, North Ryde, NSW, Australia	pmol/L	-
	Sex hormone binding globulin (138) *(sex steroid hormone pathway)*	Plasma	Electrochemiluminescence immunoassay, Elecsys 2010 analyzer, Roche Diagnostics GmbH, Mannheim, Germany	nmol/L	Total CV at 45.0 nmol/L 7% (6% within and 4% between) ICC** 0.90 (95%CI, 0.98 – 0.95)
	Estradiol (138) *(sex steroid hormone pathway)*	Plasma	Electrochemiluminescence immunoassay, Elecsys 2010 analyzer, Roche Diagnostics GmbH, Mannheim, Germany	pmol/L	Total CV at 157 pmol/L 10% (8% within and 6% between) ICC*** -

continued

Table 4-3 continued Available biomarker measurements in each study that were considered as potential mediators of the effect of adiposity on cancer risk

Cohort/ related chapter in thesis	Biomarker (ref) *(pathway)*	Measured in	Analysis method and platform, and manufacture	Unit	CV/ICC
EPIC/ Chapter 7	Adiponectin (140) *(inflammatory status pathway)*	Plasma	Enzyme-linked immunosorbent assay, R&D Systems Europe	µg/mL	CV at 4.3 µg/mL 8.2% within 7.7% between CV at 4.9 µg/mL 5.6% within 7.9% between CV at 14.7 µg/mL 5.1% within 7.6% between
	Interleukin-6 (77) *(inflammatory status pathway)*	Serum (EDTA plasma for Swedish participants)	Enzyme-linked immunoassays, R&D Systems, Europe	pg/mL	CV derived from quality control samples 6.3% within 8.2% between
	Interleukin-1 receptor antagonist (77) *(inflammatory status pathway)*	Serum (EDTA plasma for Swedish participants)	Beadbased Immunoassay, Linco, Millipore, Billerica, MA, USA	pg/mL	CV derived from quality control samples 15% within 27.7% between
	C-reactive protein (77) *(inflammatory status pathway)*	Serum (EDTA plasma for Swedish participants)	Enzyme-linked immunoassays, R&D Systems, Europe	ng/mL	CV derived from quality control samples 6.8% within 9.3% between
	Tumour necrosis factor-a (78) *(inflammatory status pathway)*	Serum	Enzyme-linked immunosorbent assays, R&D Systems, Europe	pg/mL	CV 15.0% within 20.0% between
	Tumour necrosis factor receptor-1 (78) *(inflammatory status pathway)*	Plasma	Enzyme-linked immunosorbent assays, R&D Systems, Europe	pg/mL	CV <8% within and between
	Tumour necrosis factor receptor-2 (78) *(inflammatory status pathway)*	Plasma	Enzyme-linked immunosorbent assays, R&D Systems, Europe	pg/mL	CV <8% within and between
	C-peptide (94) *(insulin signalling pathway)*	Serum	Immunoradiometric assay, Immunotech, Marseille, France	nmol/L	CV at 0.50 nmol/L 2.6% within 11.3% between
	Sex hormone binding globulin (113) *(sex steroid hormone pathway)*	Serum	Solid phase sandwich IRMA, Cis-Bio International, Gif-sur-Yvette, France	nmol/L	CV at 30 nmol/L 3.9% within 11% between
	Estradiol (113) *(sex steroid hormone pathway)*	Serum	Radio immunoassay with a double antibody system, Diagnostic Systems Laboratories Inc., Webster, TX, USA	pmol/l	CV at 294 pg/mL 8.4% within 20% between
	Estrone (113) *(sex steroid hormone pathway)*	Serum	Radio immunoassay with a double antibody System, Diagnostic Systems Laboratories Inc., Webster, TX, USA	pg/mL	CV at 55.4 pg/mL 7.4% within 16% between

continued

Table 4-3 continued Available biomarker measurements in each study that were considered as potential mediators of the effect of adiposity on cancer risk

Cohort/ related chapter in thesis	Biomarker (ref) *(pathway)*	Measured in	Analysis method and platform, and manufacture	Unit	CV/ICC
WHI-OS/ Chapter 8	Adiponectin (59, 68, 76) *(inflammatory status pathway)*	EDTA plasma	Milliplex Human Adipokine Panel-A, EMD Millipore, Billerica, MA, USA	µg/mL	CV 13% between, within not reported
	Leptin (59, 68, 76) *(inflammatory status pathway)*	EDTA plasma	Milliplex Human Adipokine Panel-B, EMD Millipore, Billerica, MA, USA	ng/mL	CV 9% between, within not reported
	Resistin (59, 68, 76) *(inflammatory status pathway)*	EDTA plasma	Milliplex Human Adipokine Panel-A, EMD Millipore, Billerica, MA, USA	ng/mL	CV 12% between, within not reported
	Hepatocyte growth factor (59, 68, 76) *(inflammatory status pathway)*	EDTA plasma	Milliplex Human Adipokine Panel-B, EMD Millipore, Billerica, MA, USA	pg/mL	CV 12% between, within not reported
	Plasminogen activator inhibitor (59, 68, 76) *(inflammatory status pathway)*	EDTA plasma	Milliplex Human Adipokine Panel-A, EMD Millipore, Billerica, MA	ng/mL	CV 12% between, within not reported
	Tumour necrosis factor-a (59, 68, 76) *(inflammatory status pathway)*	EDTA plasma	Milliplex Human Adipokine Panel-B, EMD Millipore, Billerica, MA, USA	pg/mL	CV 13% between, within not reported
	Interleukin-6 (59, 68, 76) *(inflammatory status pathway)*	EDTA plasma	Ultrasensitive solid-phase sandwich ELISA, R&D Systems, Minneapolis, MN, USA	pg/mL	CV 10% between, within not reported
	C-reactive protein (59, 68, 76) *(inflammatory status pathway)*	EDTA plasma	Latex-enhanced immunonephelometry on the Behring nephelometer II analyzer, Behring Diagnostics, San Jose, CA, USA	µg/mL	CV 4% between, within not reported
	Fasting insulin (83, 90, 119) *(insulin signalling pathway)*	Serum	Medical Research Laboratories, Highland Heights, KY, USA using assay with sensitivity of 0.26 µIU/mL	uIU/mL	Tests were performed in duplicates and those with CV >10% were repeated
	Estradiol (83, 90, 119) *(sex steroid hormone pathway)*	Serum	Vitros-Eci Immunodiagnostic assay, Ortho-Clinical Diagnostics, High Wycombe, UK	pg/mL	Tests were performed in duplicates and those with CV >20% were repeated

Abbreviation MCCS Melbourne Collaborative Cohort Study; EPIC European Prospective Investigation into Cancer and Nutrition; WHI-OS Women's Health Initiative Observational Study; CV coefficient of variation; ICC intraclass correlation coefficient; IQC internal quality control; CI confidence interval

*within the UK Biobank, the ICCs were calculated using data from a subgroup of participants who had measurements available for two time points (approximately four years apart), using linear mixed-effects modelling to obtain estimates of the between- and within-person variances.

within the MCCS, the ICC for SHBG was reported using data from a reliability study done on a subset of women with blood samples given at baseline and approximately one year later. * the reliability study was not performed for estradiol due to lack of sufficient samples

4.2 Statistical analysis

As previously described, this thesis made use of a cohort study (UK Biobank), two case-cohort studies (MCCS and WHI-OS) and a matched nested case-control study (EPIC). For each, a causal mediation analysis approach was used that lent itself to the design and characteristics of the study. These approaches, together with the method used to handle missing data, and the exposure, outcome, mediator, and confounder variables in each study are shown in Table 4-4. A detailed description of the statistical analysis methods appears in the relevant chapters.

In all studies, cancer diagnosis during the follow-up period was treated as binary outcome, because methods for handling multiple mediators with time to event outcomes are not yet fully developed (144). The two case-cohort studies were thus analysed as unmatched case-control studies with cumulative sampling, after excluding participants who died or were lost to follow up. Similarly, for the UK Biobank, participants who died or were lost to follow up were excluded.

In all primary mediation analyses, the adiposity measures were treated as binary rather than continuous. This choice was made firstly to assist with the interpretations of the mediated effects and translation to policy. Additionally, the validity of the mediation analysis using adiposity measures as continuous variables requires the assumption that the indirect and direct effects are the same for any increment in exposure, regardless of the starting value, which might be unlikely.

In the four studies, the mediation analyses aimed to quantify the mediating role of multiple interrelated pathways, as captured by one or several potentially correlated biomarkers, in the effect of adiposity on cancer risk. At least two approaches, namely

the sequential analysis for estimation of average natural indirect (NIE) and direct (NDE) effects (described in **3.7.2 Sequential analysis** (129)) and the interventional analysis for estimation of interventional indirect (IIE) and direct effects (IDE) (described in **3.7.3 Interventional analysis** (130)) allow such investigation.

Contrary to the sequential analysis, interventional effects can be used without assuming a causal ordering between the mediators (Table 3-1) (130). This is a considerable advantage for studies exploring biological pathways as potential mediators of the effect of adiposity on cancer risk. As reviewed in **2-3 Mechanisms linking adiposity to cancer**, despite the existing evidence about the possible dependence of some pathways on the others, the interrelation between pathways are complex and not entirely understood. Therefore, the interventional analysis was highly attractive and used for the analysis of the MCCS and WHI-OS case-cohort studies.

The interventional analysis can be computationally intensive because it relies on averaging estimates based on a large number of Monte-Carlo draws from the mediator distributions (Table 3-1) (130). Therefore, for the analysis of the UK Biobank data, where the analytic datasets included over 130,000 women and 170,000 men, instead, sequential analysis using the weighting approach (described **3.7.2 Sequential analysis**) was employed. The weighting method was chosen over the regression-based approach because it required fewer models to be fitted and was able to handle possible exposure-mediator or mediator-mediator interactions (Table 3-1) (129).

It was also not possible to use the interventional approach for the analysis of the matched case-control study nested within the EPIC because this approach relied on estimating absolute risks (Table 3-1) (130). Although a weighting scheme based on the inverse of

the inclusion probability of controls in the nested case-control study could have been used to estimate the absolute risks (145), using weights concurrently with the Monte-Carlo simulation introduced complications in getting the nonparametric bootstrap confidence intervals. Therefore, for this study, the regression-based sequential mediation analysis (described in **3.7.2 Sequential analysis**) was employed. The weighting method would have been preferred since it was able to handle possible exposure-mediator or mediator-mediator interactions but this approach again has not been adapted for a nested case-control study (129). For both sequential analyses, the assumed causal ordering between the pathways was based on the reviewed evidence (described in **2-3 Mechanisms linking adiposity to cancer** and illustrated in Figure 2-2).

In the analyses of studies where >10% of the eligible participants had missing values for biomarker measurements (the UK Biobank, MCCS, and EPIC), missing biomarker data were handled using multiple imputation (MI) based on chained equations. Details of MI procedures are described in the relevant chapters.

In all analyses, the 95% confidence intervals (CIs) around the indirect and direct effects were obtained using at least 1,000 bootstrap samples. In the studies where MI was used to handle missing data, mediation analysis and bootstrapping were performed for each imputed dataset. The estimated effects were then pooled across the datasets to calculate the final MI estimates of effects and the standard errors were derived from the within- and between-imputation variances based on Rubin's rules (146).

Table 4-4 Overview of the statistical analyses employed for each study

Cohort/ related chapter in thesis	Study design	Mediation analysis approach	% with missing biomarker data and approach to handle missing data	Outcome	Exposure	Mediators: Inflammatory status pathway	Mediators: Insulin signalling pathway	Mediators: Sex-steroid hormone pathway	Baseline confounders included in analysis*	N participants in the analytic dataset for mediation analysis
UK Biobank Chapter 5	Cohort	Sequential analysis using a weighting approach to estimate the natural indirect and direct effects	**Men:** 18% **Women:** 35% Multiple imputation	Colorectal cancer in men and in post-menopausal women	WC, BMI, WHR	CRP	HbA1c	SHBG Testosterone	Baseline age, education, Townsend deprivation index, family history of CRC, physical activity, alcohol intake, smoking status, red or processed meat intake, aspirin or ibuprofen use *Additionally, for women*: number of live births, previous use of oral contraceptive pill or hormone therapy	**Men:** 1,280 cancer affected & 178,656 unaffected **Women:** 790 cancer affected & 135,788 unaffected
MCCS Chapter 6	Case-cohort study treated as unmatched case-control study with cumulative sampling	Interventional analysis using a regression-standardisation approach to estimate the interventional indirect and direct effects	37% Multiple imputation	ER-positive breast cancer in post-menopausal women	BMI, WC	-	Fasting insulin	Free estradiol (calculated from measured SHBG and total estradiol)	Baseline age, country of birth, socioeconomic position, physical activity, alcohol intake, smoking status, history of cardiovascular disease or arthritis, age at menarche, previous use of oral contraceptive pill or hormone therapy, parity and age at first live birth, lactation duration	149 cases & 1,029 controls

continued

Table 4-4 continued Overview of the statistical analyses employed for each study

Cohort/ related chapter in thesis	Study design	Mediation analysis approach	% with missing biomarker data and approach to handle missing data	Outcome	Exposure	Mediators Inflammatory status pathway	Insulin signalling pathway	Sex-steroid hormone pathway	Baseline confounders included in analysis	N participants in the analytic dataset for mediation analysis
EPIC Chapter 7	Matched nested case-control, but matching was broken in all analyses	Sequential analysis using a regression-based approach to estimate the natural indirect and direct effects	17% Multiple imputation analyses	Endometrial cancer in post-menopausal women	BMI, WC	Adiponectin IL-6 IL-1Ra TNF-R1 TNF-R2 CRP	C-peptide	Estrone and Free estradiol (calculated from measured SHBG and total estradiol)	Baseline age, country region, physical activity, and smoking status, age at menarche, number of full-term pregnancies, previous use of oral contraceptive pill or hormone therapy	163 cases & 306 control
WHI-OS Chapter 8	Case-cohort study treated as unmatched case-control study with cumulative sampling	Interventional analysis using a regression-standardisation approach to estimate the interventional indirect and direct effects	7% in all three case-control studies Complete case analysis	ER-positive breast, endometrial, and colorectal cancer in post-menopausal women	WC, BMI, WHR	Adiponectin Leptin Resistin HGF PAI IL-6 CRP	Fasting insulin	Total estradiol	Baseline age, education, physical activity, alcohol intake, ever smoked, nonsteroidal anti-inflammatory drug use, age at menarche, parity, previous use of oral contraceptive pill or hormone therapy	*Breast cancer:* 187 cases & 311 controls *Endometrial cancer:* 98 cases & 214 controls *Colorectal cancer:* 193 cases & 288 controls

Abbreviations MCCS Melbourne Collaborative Cohort Study; EPIC European Prospective Investigation into Cancer and Nutrition; WHI-OS Women's Health Initiative Observational Study; ER estrogen receptor; WC waist circumference; BMI body mass index; WHR waist-hip ratio; CRP C reactive protein; IL interleukin; IL-1RA interleukin-1 receptor antagonist; TNF tumour necrosis factor; TNF-R tumour necrosis factor receptor; HGF hepatocyte growth factor; PAI plasminogen activator inhibitor; HbA1c Glycosylated haemoglobin; SHBG sex hormone binding globulin

*Potential confounders (common causes of exposure-outcome, mediator-outcome, or exposure-mediator association) were identified *a priori* based on existing evidence as well as expert opinion

Chapter 5: Adiposity and colorectal cancer risk in men and postmenopausal women in the UK Biobank

In this chapter, an analysis of the UK Biobank cohort has been carried out to investigate the mediating roles of C-reactive protein (CRP, to reflect the inflammatory status), glycosylated haemoglobin (HbA1c, to reflect insulin resistance), and sex hormone-binding globulin (SHBG) and testosterone (to reflect alterations in sex-steroid hormone levels) in the effect of adiposity on colorectal cancer risk in men and postmenopausal women. This chapter includes the author accepted version of the following paper, which is in production for publication at the time of thesis submission:

Dashti SG, Viallon V, Simpson JA, Karahalios A, Moreno-Betancur M, English DR, Gunter MJ, Murphy N. Explaining the link between adiposity and colorectal cancer risk in men and postmenopausal women in the UK Biobank: a sequential causal mediation analysis. Int J Cancer. (1097-0215 (Electronic)). DOI: 10.1002/ijc.32980

For consistency of thesis chapters, spelling, referencing, tables, and figure numbers have been modified. Supplementary material related to this publication is presented in Appendix A.

Explaining the link between adiposity and colorectal cancer risk in men and postmenopausal women in the UK Biobank: a sequential causal mediation analysis

S Ghazaleh Dashti[1,2], Vivian Viallon[2], Julie A Simpson[1], Amalia Karahalios[1,3], Margarita Moreno-Betancur[4,5], Dallas R English[1,6], Marc J Gunter[2], Neil Murphy[2]

1 Centre for Epidemiology and Biostatistics, Melbourne School of Population and Global Health, The University of Melbourne, Melbourne, Victoria, Australia
2 Section of Nutrition and Metabolism, International Agency for Research on Cancer (IARC), Lyon, France
3 School of Public Health and Preventive Medicine, Monash University, Melbourne, Victoria, Australia
4 Clinical Epidemiology and Biostatistics Unit, Murdoch Children's Research Institute, Melbourne, Australia
5 Department of Paediatrics, The University of Melbourne, Melbourne, Victoria, Australia
6 Cancer Epidemiology and Intelligence Division, Cancer Council Victoria, Melbourne, Victoria, Australia

Corresponding author
S Ghazaleh Dashti
Centre for Epidemiology and Biostatistics
Melbourne School of Population and Global Health
Level 3, 207 Bouverie Street
The University of Melbourne VIC 3010 Australia
Email: seyedeh.dashti@unimelb.edu.au

Short title: *Explaining adiposity and colorectal cancer link*

Key words: adiposity, colorectal cancer, causal mediation analysis, C-reactive protein, inflammation, haemoglobin-A1c, insulin resistance, sex-steroid hormones, sex hormone-binding globulin, testosterone

Article category: Cancer Epidemiology

Abbreviations used: CRC: colorectal cancer; CRP: C-reactive protein; HbA1c: haemoglobin-A1c; SHBG: sex hormone binding globulin; NDE: natural direct effect; NIE: natural indirect effect; RR: risk ratio; CI: confidence interval; ICD: International Classification of Disease; IQC: internal quality control; CV: coefficient of variation; WHO: World Health Organization; GMR: geometric mean ratio; OR: odds ratio; TE: total effect; IPW: inverse probability weights; MI: multiple imputation; HPFS: Health Professionals Follow-up Study; sOB-R: soluble leptin receptor; HR: hazard ratio; TyG: triglyceride-glucose index

Novelty and Impact: Mechanisms underlying adiposity-colorectal cancer (CRC) association are not well understood. Here, we used UK Biobank data and performed sequential mediation analysis to quantify mediating effects of C-reactive protein,

haemoglobin-A1c, sex hormone binding globulin and testosterone in this association in men and postmenopausal women. Our results suggest pathways marked by these obesity-related factors may not explain a large proportion of this association and pathways other than those captured here may also be important for understanding the adiposity-CRC link.

Abstract

Mechanisms underlying adiposity-colorectal cancer (CRC) association are incompletely understood. Using UK Biobank data, we investigated the role of C-reactive protein (CRP), haemoglobin-A1c (HbA1c), and (jointly) sex hormone-binding globulin (SHBG) and testosterone, in explaining this association. Total effect of obesity vs normal-weight (based on waist circumference, body mass index, waist-hip ratio) on CRC risk was decomposed into natural direct (NDE) and indirect (NIE) effects using sequential mediation analysis. After a median follow-up of 7.1 years, 2070 incident CRC cases (men=1280; postmenopausal women=790) were recorded. For men, the adjusted risk ratio (RR) for waist circumference (≥102 *vs* ≤94cm) was 1.37 (95% confidence interval (CI),1.19-1.58). The RRs^{NIE} were 1.08 (95%CI,1.01-1.16) through all biomarkers, 1.06 (95%CI,1.01-1.11) through pathways influenced by CRP, 0.99 (95%CI,0.97-1.01) through HbA1c beyond [the potential influence of] CRP, and 1.03 (95%CI,0.99-1.08) through SHBG and testosterone combined beyond CRP and HbA1c. The RR^{NDE} was 1.26 (95%CI,1.09-1.47). For women, the RR for waist circumference (≥88 *vs* ≤80cm) was 1.27 (95%CI,1.07-1.50). The RRs^{NIE} were 1.08 (95%CI,0.94-1.22) through all biomarkers, 1.08 (95%CI,0.99-1.17) through CRP, 1.00 (95%CI,0.98-1.02) through HbA1c beyond CRP, and 1.00 (95%CI;0.92-1.09) through SHBG and testosterone combined beyond CRP and HbA1c. The RR^{NDE} was 1.18 (95%CI,0.96-1.45). For men and women, pathways influenced by CRP explained a small proportion of the adiposity-CRC association. Testosterone and SHBG also explained a small proportion of this association in men. These results suggest that pathways marked by these obesity-related factors may not explain a large proportion of the adiposity-CRC association.

Introduction

Globally, colorectal cancer (CRC) is the third most commonly diagnosed cancer among men and women (31). There is a wide disparity in CRC incidence across the world regions. High-income countries have three times higher incidence rate compared with low- and middle-income countries (31), while the latter, especially countries undergoing significant developments, have faced an increased burden of CRC more recently (32). These patterns highlight the role that some of the consequences of a 'Westernized' lifestyle may have in developing CRC (31).

Excess body fatness, a hallmark of 'Westernized' lifestyle, is an established risk factor for CRC, as well as for 12 other cancers (30). Based on a recent meta-analysis, the relative risk (RR) of CRC for the highest *vs* lowest category of waist circumference (a measure of central adiposity) was 1.48 (95% confidence interval (CI), 1.30-1.67) for men and 1.44 (95%CI, 1.30-1.60) for women (33). Similar associations were reported for BMI (obese *vs* normal), with a more evident difference in the strength of association for men (RR 1.47; 95%CI, 1.36-1.58) compared with women (RR 1.15; 95%CI, 1.08-1.23) (33).

Overweight and obesity already make a substantial contribution to the burden of CRC. In 2012, 13% (90% confidence interval (CI), 11%-14%) of colon and 6% (90%CI, 5%-7%) of rectal cancer cases diagnosed in men and 8% (90%CI, 7%-8%) of colon and 4% (90%CI, 3%-4%) of rectal cancer cases diagnosed in women were attributable to overweight and obesity (BMI ≥25 kg/m^2) (3). Globally, the prevalence of adults living with obesity (BMI >30 kg/m^2) has increased from ~3% in 1975 to ~11% in 2014 in men and from ~6% in 1975 to ~15% in 2014 in women (28). If these trends continue, by 2025, 18% of men and 21% of women will have obesity (4). It can, therefore, be expected for obesity to make considerably larger contributions to CRC burden in the future.

Adiposity-induced chronic inflammation, dysregulation of insulin signalling and glucose homeostasis, and sex-steroid hormones production are three interconnected pathways hypothesized to mediate the adiposity–CRC association (147). Inflammation might exert its influence on CRC risk directly (through its angiogenic, mitogenic, and antiapoptotic effects) (147) and through contributions to the development of insulin resistance and dysregulated insulin signalling (8) and sex-steroid hormones production pathways (10, 148). Insulin resistance may increase cancer risk through hyperglycaemia and exposure to elevated levels of insulin which has mitogenic and antiapoptotic effects (147) and its influences on insulin-like growth factor-1 (14, 15), and sex-steroid hormones levels (10, 148, 149). The potential role of sex-steroid hormones in colorectal carcinogenesis is not well-understood, but the higher incidence of CRC in men, a stronger adiposity-CRC association in men (147), differences in the associations between adiposity and sex-steroid hormone levels in men and women (inverse association with testosterone levels in men, positive association in women) (150), and the inverse association between hormone therapy and CRC in women (100) are consistent with a role of sex-steroid hormones in the aetiology of CRC (147). Overall, however, the evidence for associations between sex-steroid hormones and CRC risk in men and women is limited and inconclusive (147).

We used data from the UK Biobank to investigate the extent to which biomarkers of these three pathways mediate the effect of adiposity (measured by waist circumference, BMI, and waist-hip ratio) on CRC risk in men and postmenopausal women. A sequential causal mediation analysis approach was used to quantify the mediating effects through the three pathways without assuming they acted independently (129). The assessed biomarkers were high sensitivity C-reactive protein (CRP, a non-specific biomarker of systemic inflammation), glycated haemoglobin (HbA1c, a biomarker of glycaemic control and a

correlate of insulin resistance (82)), and sex hormone binding globulin (SHBG) and testosterone (treated jointly to capture some aspects of altered sex-steroid hormone levels in the obese state).

Methods

The UK Biobank is a prospective cohort of 229,134 men and 273,402 women recruited between 2006 and 2010 following an invitation letter sent to about 9.2 million individuals (~5.5% response rate) (136). Invitees were registered with the UK National Health Service, were mostly 40-69 years old, and lived within approximately 25 miles of one of 22 assessment centres in England, Wales, and Scotland. Participants gave informed consent, and 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 approved the study. In 2004, the Ethics and Governance Council was established to ensure adherence to the Ethics and Governance Framework (http://www.ukbiobank.ac.uk/ethics/) of the study. This research was conducted using the UK Biobank Resource under application number 25897.

The current study used data from all participants, with the outcome defined as CRC diagnosis between one year after baseline (i.e. recruitment) and end of follow-up (see *Follow-up and outcome assessment*).

Data collection

At baseline, information on demographics, early life and lifestyle exposures, health history and medication use were collected using a self-administered touchscreen questionnaire available at http://www.ukbiobank.ac.uk/key-documents/. Weight, height, waist and hip circumference were measured following prespecified protocols (151), and

blood samples were collected, centrifuged, and stored at -80°C or liquid nitrogen within ~24 hours after venepuncture (152).

Follow-up and outcome assessment

Incident cancer cases and cancer cases recorded first in death certificates were identified through linkage to national cancer and death registries. Complete follow-up was available through 31 March 2016 for England and Wales and 31 October 2015 for Scotland. The 10th Revision of the International Classification of Disease (ICD-10) was used to define CRC cases (C18, C19, and C20 for malignant neoplasms of colon, rectosigmoid junction, and rectum respectively).

Selection of participants eligible for the study

Participants were ineligible for this study if they had withdrawn consent; had cancer diagnosed before baseline; had less than one-year follow-up; were diagnosed with CRC within the first year after baseline; were diagnosed with cancer of the appendix (ICD-10 C18.1; <1%); died (<2%) or were lost to follow-up (<0.5%) before the end of follow-up; did not have body size measurements collected or had BMI<18.5 kg/m^2 at baseline; had unknown diabetes status, had diabetes or reported taking diabetes medication at baseline (Figure 5. 1 and Figure 5. 2). The last exclusion criterium was imposed because diabetes and diabetes medication use may change the link between adiposity and insulin signalling (153). Additionally, women who, at baseline, were premenopausal or ≤55 years old with unknown menopause status, who reported taking hormone therapy, or hormonal contraceptives were ineligible (Figure 5. 2). We did not include premenopausal women in this analysis because the mechanistic pathways, especially the sex-steroid hormones pathway, might be different in pre- and postmenopausal women. Instead of multiply imputing missing confounder data (the approach used to handle missing biomarker data),

we also excluded participants with missing data for any of the selected confounders (see *Confounder selection*) because they were a relatively small proportion (≤10%) of all the participants.

Figure 5-1 Flow diagram demonstrating the selection process of men in the UK Biobank for the present study

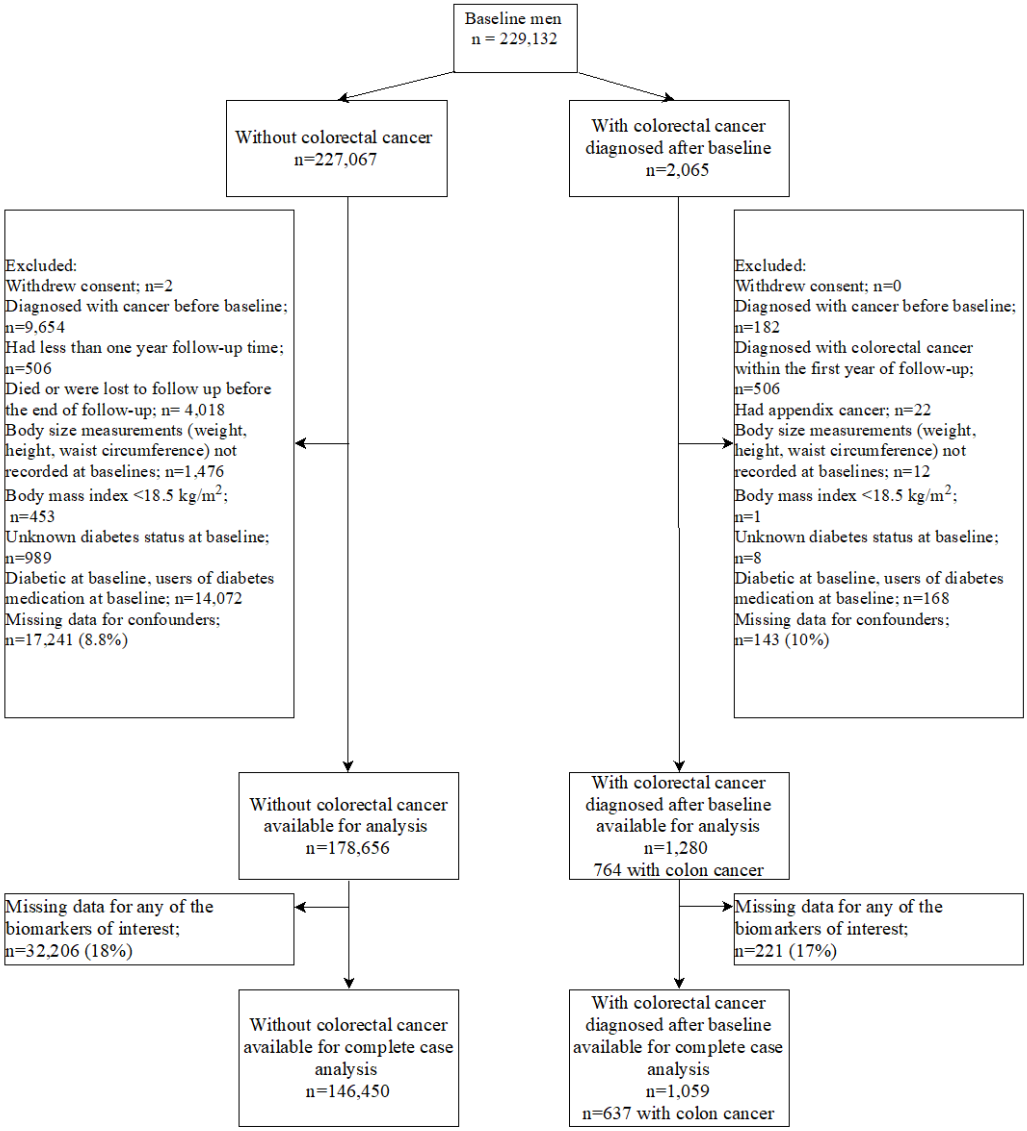
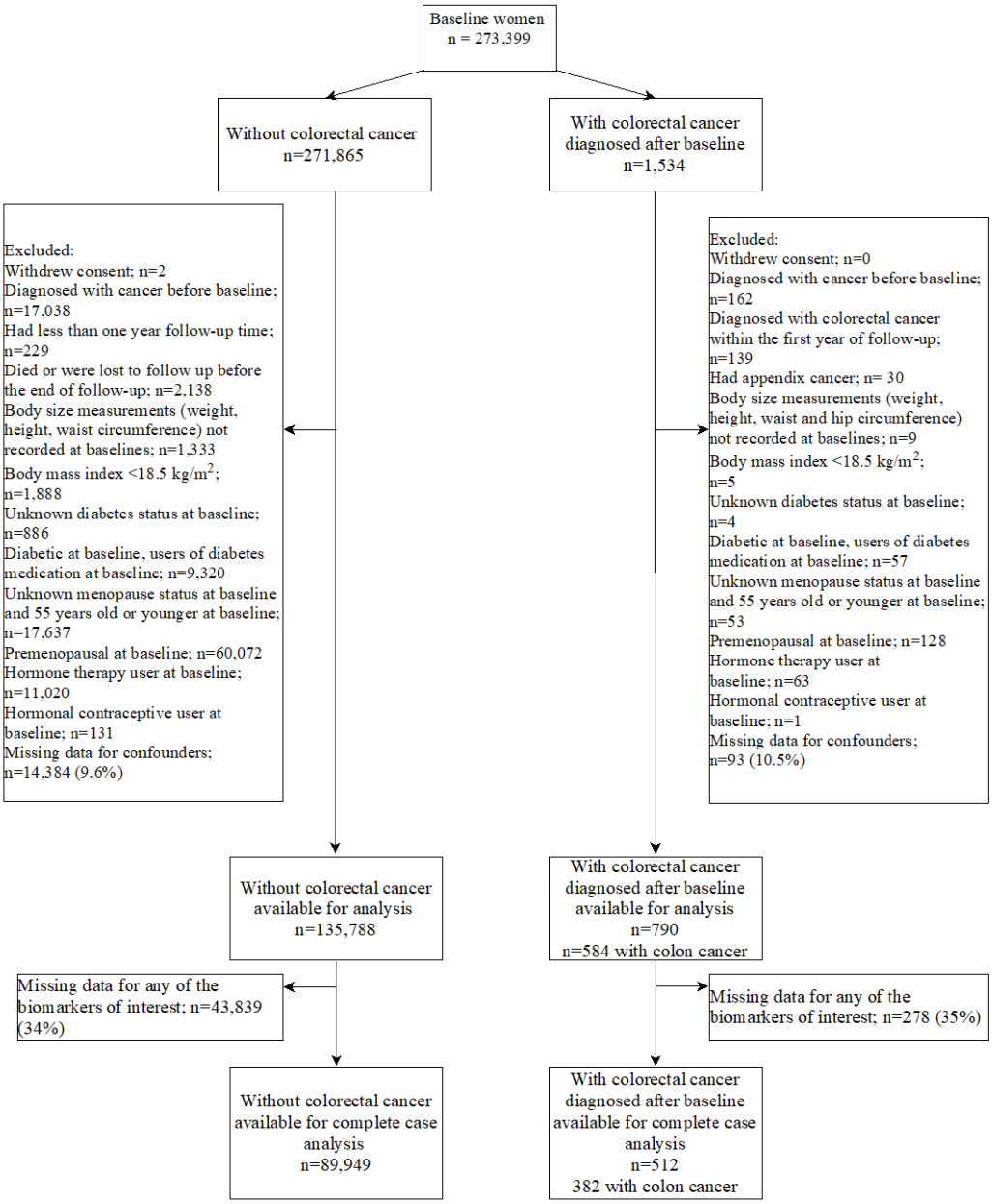


Figure 5-2 Flow diagram demonstrating the selection process of women in the UK Biobank for the present study



Biomarkers and laboratory methods

Descriptions of laboratory methods and protocols for measuring the biomarkers have been published (142, 143). We did not include estradiol in our analyses because most participants had poor quality data (142). Measurements were performed in blood samples (red blood cells for HbA1c, serum for all other biomarkers) collected at baseline (2006-2010) as part of the UK Biobank Biomarker Project. Immuno-turbidimetric metric (Beckman Coulter AU5800) was used to measure CRP (with average within-laboratory coefficient of variation (CV) of 2.3% for low and 1.7% for high internal quality control (IQC) samples); high-performance liquid chromatography (Bio-Rad Variant II Turbo) to measure HbA1c (CV 2.1% for low, 1.5% for high IQC samples); and chemiluminescent immunoassay (Beckman Coulter DXI 800) to measure SHBG (CV 5.7% for low, 5.2% for high IQC samples), and testosterone (CV 8.3% for low, 4.2% for high IQC samples) (142, 143). An algorithm was developed to ensure each laboratory batch included a random selection of participants to reduce the likelihood of any systematic differences in assay results between participants (142). When applicable, the data we received had also been corrected for aliquot-dilution effect (SHBG, testosterone) and date of the assay (HbA1c, SHBG, testosterone) (152). Free testosterone was calculated from measured testosterone, SHBG, and albumin (154). For a subgroup of participants, biomarker measurements were additionally performed in blood samples collected at repeat assessment visit (2012-2013). Intraclass correlation coefficients (ICC) were calculated using the two measurements to assess the reproducibility of the measures (see Statistical analysis).

Statistical analysis

Waist circumference, BMI, and waist-hip ratio were used as measures of adiposity in separate analyses. For BMI, we used the World Health Organisation (WHO)

recommended cut-off values to compare participants with obesity *vs* normal weight (≥30 *vs* <25 kg/m^2). For waist circumference, we also used the WHO suggested cut-off values to compare participants defined as "at substantially increased risk of metabolic complications" (referred to as with obesity hereafter) with normal (>102 *vs* ≤94 cm for men, >88 *vs* ≤80 cm for women) (25). For waist-hip ratio, we used sex-specific quartiles to define with obesity (4th quartile >0.98 for men, >0.86 for women) *vs* normal (1st quartile ≤0.89 for men, ≤0.77 for women) because the WHO does not have a suggested cut-off value for normal (25). We treated CRC diagnosis as a binary outcome because methods for handling multiple mediators and survival outcome are not fully developed yet (144).

To calculate the ICCs for biomarker measurements, linear mixed-effects modelling was used to obtain estimates of the between- and within-person variances.

Confounder Selection: The following variables were identified *a priori* and included as potential confounders (i.e. common causes for exposure-outcome, mediator-outcome, or exposure-mediator associations (155)): age at baseline (fitted as splines), educational attainment, Townsend deprivation index, first-degree relative history of CRC, physical activity, alcohol intake, smoking status, total red or processed meat intake, regular aspirin or ibuprofen use; additionally for women number of live births, ever use of oral contraceptive pill, and ever use of hormone therapy.

Adiposity-biomarker associations: For each biomarker, the geometric mean ratio (GMR; 95% CI) was estimated in relation to adiposity measures (comparing with obesity *vs* normal) using linear regression models fitted to the log-transformed biomarker variable adjusting for potential confounders.

Adiposity-CRC & Biomarker-CRC associations: Odds ratios (OR; 95% CI) for the adiposity-CRC and biomarker-CRC associations were estimated using logistic regression models adjusted for confounders. For each biomarker-CRC model, we additionally included waist circumference and other biomarkers that might have confounded the association (i.e. CRP for HbaA1c-CRC model; and CRP and HbA1c for SHBG-CRC and testosterone-CRC models, selected based on the hypothesized causal sequence – see Figure 5. 3). The linearity of the biomarker-CRC associations was investigated in models with biomarker variables fitted as restricted cubic splines (2 degrees of freedom).

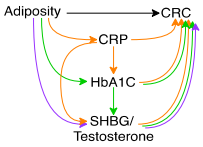
Sequential mediation analysis: Using a causal mediation analysis approach (129), the total adiposity-CRC association (total effect Risk Ratio $(RR)^{TE}$) was decomposed into a natural indirect effect (RR^{NIE}; effect explained by all the included biomarkers, i.e. CRP, HbA1c, SHBG, and testosterone) and a natural direct effect (RR^{NDE}; effect not explained by any of the included biomarkers) (see (155) for a formal definition of NDE and NIE). Assuming that CRP levels might have also influenced HbA1c or biomarker of sex-steroid hormone pathway (8, 10), and HbA1c might have influenced biomarker of sex-steroid hormone pathway (10, 149), but not vice versa, we took a sequential approach to further decompose the estimated NIE into the following NIEs: i) through pathways influenced by CRP; ii) through HbA1c excluding the possible influence of CRP on HbA1c; and iii) through SHBG and testosterone combined excluding the influences of CRP or HbA1c. Figure 5. 3. provides a schematic explanation of these effects and the sequential process used to estimate them (129).

The mediation analysis involved calculating inverse probability weights (IPWs) based on the predicted probabilities of being obese from logistic regression models for exposure conditional on confounders (129) . Then, under different counterfactual scenarios,

potential outcomes weighted by the IPWs (to take confounding into account) were estimated based on logistic regression models for the outcome conditional on exposure, confounders, with and without the biomarkers (129). Potential outcomes were expressed as risks, using the following formulae $risk = \frac{1}{(1+\exp(-\log(odds)))}$ *(156).* To accommodate non-linear biomarker-CRC associations, relevant biomarker variables were included as restricted cubic splines. Outcome models used to estimate each effect are presented in Figure 5. 3.

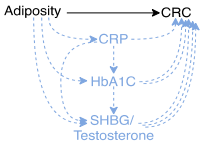
Figure 5-3 Schematic interpretation of the estimated total, natural direct, and natural indirect effects In this study, C-reactive protein (CRP) was used as a general marker for inflammation, HbA1C as a marker for glycaemic control and a correlate of insulin resistance, and sex hormone binding globulin (SHBG) and testosterone treated jointly as makers of the sex-steroid hormone pathway. Waist circumference, BMI, and waist-hip ratio were used as proxy measures for adiposity in separate sets of analyses. In each provided outcome model needed to estimate the effects, y is the outcome (colorectal cancer incidence), a is the exposure (each adiposity measure comparing obese vs normal), c confounders, m1 CRP, m2 HbA1c, m3a SHBG, and m3b testosterone.

The total effect of adiposity and CRC
This effect captures the adiposity-CRC association through the included biomarkers (the indirect effect) and not through the included biomarkers (referred to as the direct effect).



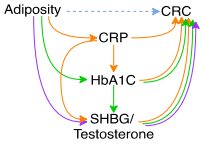
Required models for the outcome — *logistic y a c*

The natural direct effect (NDE) not through the included biomarkers
This effect captures the adiposity-CRC association not through any of the included biomarkers (the solid arrow).



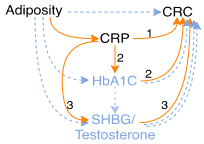
Required models for the outcome — *logistic y a c m1 m2 m3a m3b & logistic y a c*

The natural indirect effect through all included biomarkers
This effect captures the adiposity-CRC association through all the included biomarkers (the solid arrows).



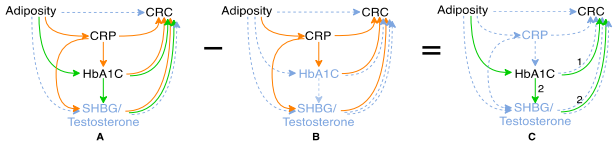
The natural indirect effect through pathways originating with CRP *Required models for the outcome* — *logistic y a c m1 m2 m3a m3b & logistic y a c*

This effect captures the effect of adiposity-induced changes in CRP levels on CRC risk directly (1), or indirectly through influencing HbA1C (2) or SHBG and testosterone (3) levels.



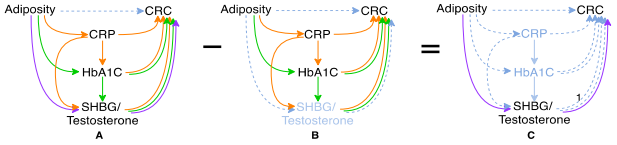
Required models for the outcome — *logistic y a c m1 & logistic y a c*

The natural indirect effect through HbA1C beyond the influence of C-reactive protein
This effect (C) is estimated as the difference between the indirect effect through CRP and HbA1C (A) and the indirect effect through CRP (B). It thus excludes the possible influence of CRP on HbA1C levels and captures the effect of
adiposity-induced changes in HbA1C on CRC risk directly (1), or indirectly through influencing SHBG and testosterone levels (2)



Required models for the outcome — *logistic y a c m1 m2 & logistic y a c* — *logistic y a c m1 & logistic y a c*

The natural indirect effect through sex-hormone-binding globulin and testosterone beyond the influences of C-reactive protein and HbA1C
This effect (C) is estimated as the difference between the indirect effect through CRP, HbA1C, and SHBG and testosterone (A) and the indirect effect through CRP and HbA1C (B). It thus excludes the possible influence of CRP or HbA1C on SHBG and testosterone levels and captures the effect adiposity-induced changes in SHBG and testosterone have on CRC risk (1).



Required models for the outcome — *logistic y a c m1 m2 m3a m3b & logistic y a c* — *logistic y a c m1 m2 & logistic y a c*

Missing data: Missing biomarker data were handled using multiple imputation (MI) based on chained equations with 20 iterations for men and 35 for women (18% of men and 35% of women had missing values for at least one biomarker (Table 5-1)). For each adiposity measure, MI was performed with separate imputation models for obese and non-obese individuals (157). The imputation models included all variables in the mediation analyses and the biomarker-biomarker and biomarker-CRC interaction terms (i.e. treating these interactions as 'just another variable' with missing values) (157). Then, the obesity-biomarker interaction terms were generated from the multiply imputed datasets (157).

Mediation analysis was performed for each imputed dataset and standard errors for TE, NDE, and NIEs were estimated using 1000 bootstrap samples. Rubin's rules were then applied to calculate the final MI estimates of effects and 95% CIs by pooling the estimated effects across the imputed datasets and deriving the standard errors from the within- and between-imputation variances (146).

Additional Analyses: The following additional mediation analyses were performed: i) for the outcome models, including an interaction term between the adiposity measure and each biomarker, as well as between CRP-HbA1c, CRP-SHBG, CRP-testosterone, HbA1C-SHBG, HbA1C-testosterone, and SHBG-testosterone; ii) using calculated free testosterone as the biomarker of sex-steroid hormone pathway (instead of testosterone and SHBG combined); iii) excluding rectal cancer cases and limiting analyses to colon cancer; and iv) excluding colorectal cancer cases diagnosed within the first two years of follow-up and non-cases with follow-up ≤2 years.

All analyses were performed using Stata version 15 (158).

Data availability

UK Biobank is an open-access resource. Bonafide researchers can apply to use the UK Biobank dataset by registering and applying at http://ukbiobank.ac.uk/register-apply/.

Results

Overall, 179,936 men and 136,578 women met the eligibility criteria for this analysis (Figure 5. 1and Figure 5. 2). Men with waist circumference >102 *vs* ≤94 cm had higher median red or processed meat intake, and a higher proportion had lower physical activity levels, were former smokers, and reported regular aspirin or ibuprofen use. Similar differences were observed for women with waist circumference >88 *vs* ≤80 (Table 5-1; see Supplementary Table A-1 and Supplementary Table A-2 for distributions of baseline characteristics by BMI and waist-hip ratio categories respectively). During a median follow-up time of 7.1 years, 1,280 and 790 CRC cases were diagnosed in men and postmenopausal women, respectively.

The ICCs (95% CI) for the biomarkers ranged from 0.56 (95%CI, 0.55-0.57) to 0.86 (95%CI, 0.86-0.87) for CRP and SHBG respectively in men and 0.67 (95%CI, 0.66-0.68) to 0.81 (95%CI, 0.80-0.82) for CRP and SHBG respectively in women (Supplementary Table A-3).

Table 5-1 Baseline Characteristics of men and women eligible for the study by categories of waist circumference

	Men			**Women**		
Waist circumference, cm	≤94 N=83,940	>94 - ≤102 N=51,228	>102 N=44,768	≤80 N=54,339	>80 - ≤88 N=35,820	>88 N=46,419
Colorectal cancer affected - yes, n (%)	478 (1)	402 (1)	400 (1)	288 (1)	174 (0)	328 (1)
Age at recruitment, median (25th - 75th percentiles)	56.0 (48.0-62.0)	58.0 (50.0-63.0)	59.0 (51.0-64.0)	60.0 (56.0-64.0)	61.0 (57.0-65.0)	61.0 (57.0-65.0)
Age at cancer diagnosis or the end of follow up, median (25th - 75th percentiles)	63.3 (55.4-69.4)	65.2 (57.5-70.5)	65.7 (58.5-70.6)	67.4 (63.2-71.0)	68.2 (64.3-71.8)	68.2 (64.4-71.8)
Follow-up time - years, median (25th - 75th percentiles)	7.1 (6.5-7.7)	7.2 (6.4-7.8)	7.1 (6.4-7.7)	7.1 (6.5-7.7)	7.1 (6.4-7.7)	7.1 (6.4-7.7)
Highest educational attainment, n (%)						
None of the below	10,106 (12)	8,040 (16)	8,674 (19)	9,026 (17)	7,609 (21)	11,596 (25)
CSEs/O-levels/GCSEs or equivalent	20,321 (24)	11,739 (23)	10,056 (22)	14,853 (27)	9,648 (27)	11,801 (25)
NVQ/HND/HNC/A-levels or equivalent	16,959 (20)	10,607 (21)	9,456 (21)	6,667 (12)	4,461 (12)	6,277 (14)
Other professional qualifications. (e.g. nurse/teacher)	24,048 (29)	14,462 (28)	11,878 (27)	17,327 (32)	10,530 (29)	12,928 (28)
College/university degree	12,506 (15)	6,380 (12)	4,704 (11)	6,466 (12)	3,572 (10)	3,817 (8)
Townsend deprivation index at recruitment - quintiles, n (%)						
1 (least degree of deprivation)	18,372 (22)	11,232 (22)	8,861 (20)	12,886 (24)	7,751 (22)	8,605 (19)
2	17,559 (21)	10,944 (21)	8,758 (20)	12,218 (22)	7,824 (22)	9,144 (20)
3	16,739 (20)	10,443 (20)	8,943 (20)	11,392 (21)	7,736 (22)	9,369 (20)
4	16,157 (19)	9,772 (19)	9,184 (21)	10,152 (19)	7,005 (20)	9,669 (21)
5 (highest degree of deprivation)	15,113 (18)	8,837 (17)	9,022 (20)	7,691 (14)	5,504 (15)	9,632 (21)
First degree relative history of colorectal cancer - yes, n (%)	8,923 (11)	6,041 (12)	5,441 (12)	6,316 (12)	4,383 (12)	5,680 (12)
Physical activity - MET hr/wk, n (%)						
<10	13,575 (16)	10,941 (21)	12,768 (29)	9,968 (18)	7,964 (22)	14,117 (30)
10-<20	13,279 (16)	9,006 (18)	8,094 (18)	9,513 (18)	6,678 (19)	8,771 (19)
20-<40	20,937 (25)	12,163 (24)	9,805 (22)	13,531 (25)	8,840 (25)	10,257 (22)
40-<60	12,273 (15)	6,546 (13)	4,739 (11)	7,739 (14)	4,530 (13)	5,088 (11)
60+	23,876 (28)	12,572 (25)	9,362 (21)	13,588 (25)	7,808 (22)	8,186 (18)
Alcohol intake, n (%)						
never	4,431 (5)	2,522 (5)	2,537 (6)	4,340 (8)	2,884 (8)	5,091 (11)
Current occasionally to 1-3 times/month	12,412 (15)	7,200 (14)	7,668 (17)	12,576 (23)	9,328 (26)	14,923 (32)
1-2 times/week	21,613 (26)	13,007 (25)	11,956 (27)	13,647 (25)	9,239 (26)	11,637 (25)
3-4 times/week	23,358 (28)	14,520 (28)	11,335 (25)	12,790 (24)	7,724 (22)	8,101 (17)
daily or almost daily	22,126 (26)	13,979 (27)	11,272 (25)	10,986 (20)	6,645 (19)	6,667 (14)
Smoking status, n (%)						
Never	47,086 (56)	24,891 (49)	19,373 (43)	33,719 (62)	20,990 (59)	25,575 (55)
Former	26,615 (32)	20,485 (40)	20,250 (45)	16,693 (31)	12,151 (34)	17,221 (37)
Current	10,239 (12)	5,852 (11)	5,145 (11)	3,927 (7)	2,679 (7)	3,623 (8)

continued

Table 5-1 continued Baseline Characteristics of men and women eligible for the study by categories of waist circumference

	Men			Women		
Waist circumference, cm	≤94	>94 - ≤102	>102	≤80	>80 - ≤88	>88
	N=83,940	N=51,228	N=44,768	N=54,339	N=35,820	N=46,419
Total red and processed meat intake - times/week, median (25th - 75th percentiles)	3.5 (2.5-5.0)	4.0 (2.5-5.0)	4.5 (2.5-5.5)	2.5 (2.0-4.0)	2.5 (2.0-4.5)	3.0 (2.0-4.5)
Regular aspirin or ibuprofen use - yes, n (%)	18,700 (22)	14,187 (28)	14,657 (33)	10,656 (20)	8,427 (24)	12,954 (28)
Number of live births, median (25th - 75th percentiles)	-	-	-	2.0 (1.0-2.0)	2.0 (1.0-3.0)	2.0 (1.0-3.0)
Ever use of oral contraceptive pill - yes, n (%)	-	-	-	43,179 (79)	28,021 (78)	35,539 (77)
Ever use of hormone therapy - yes, n (%)	-	-	-	24,233 (45)	17,395 (49)	22,551 (49)
Adiposity measures						
Waist Circumference - cm, median (25th - 75th percentiles)	88.0 (84.0-92.0)	98.0 (96.0-100.0)	108.0 (105.0-114.0)	75.0 (71.0-78.0)	84.0 (82.0-86.0)	96.0 (92.0-102.0)
Body Mass Index - kg/m2, n (%)						
≥18.5-<25	44,047 (52)	3,134 (6)	93 (0)	41,836 (77)	9,267 (26)	1,213 (3)
≥25-<30	38,951 (46)	39,700 (77)	12,529 (28)	12,243 (23)	23,690 (66)	17,688 (38)
≥30	942 (1)	8,394 (16)	32,146 (72)	260 (0)	2,863 (8)	27,518 (59)
Body Mass Index - kg/m2, median (25th - 75th percentiles)	24.9 (23.3-26.3)	27.9 (26.6-29.3)	31.6 (29.7-34.0)	23.3 (21.8-24.8)	26.4 (24.9-28.0)	30.9 (28.5-34.0)
Waist-Hip Ratio, n (%)						
≤0.89 M; ≤0.77 W	44,766 (53)	3,681 (7)	409 (1)	27,045 (50)	3,140 (9)	874 (2)
>0.89 - ≤0.98 M; >0.77 - ≤0.86 W	37,589 (45)	37,444 (73)	16,277 (36)	26,254 (48)	25,958 (72)	18,284 (39)
0.98 M; >0.86 W	1,585 (2)	10,103 (20)	28,082 (63)	1,040 (2)	6,722 (19)	27,261 (59)
Waist-Hip ratio, median (25th - 75th percentiles)	0.9 (0.9-0.9)	0.9 (0.9-1.0)	1.0 (1.0-1.0)	0.8 (0.7-0.8)	0.8 (0.8-0.9)	0.9 (0.8-0.9)
Biomarkers, median (25th - 75th percentiles)						
C-reactive protein - mg/L	0.9 (0.5-1.7)	1.3 (0.8-2.4)	2.0 (1.1-3.6)	0.8 (0.4-1.6)	1.4 (0.8-2.6)	2.5 (1.4-4.7)
C-reactive protein missing value - yes, n (%)	4,369 (5)	2,633 (5)	2,601 (6)	3,053 (6)	2,049 (6)	2,943 (6)
Glycated hemoglobin- nmol/mol	34.4 (32.1-36.6)	35.0 (32.6-37.4)	36.0 (33.4-38.6)	35.2 (33.2-37.3)	35.7 (33.6-37.9)	36.6 (34.3-39.1)
Glycated hemoglobin- missing value - yes, n (%)	4,731 (6)	2,844 (6)	2,643 (6)	3,284 (6)	2,277 (6)	3,134 (7)
Sex Hormone Binding Globulin - nmol/L	40.8 (31.3-52.4)	35.2 (27.0-45.3)	32.9 (24.9-42.6)	67.5 (52.4-85.7)	53.1 (40.3-68.8)	42.1 (31.3-55.9)
Sex Hormone Binding Globulin missing value - yes, n (%)	11,043 (13)	6,543 (13)	5,843 (13)	8,083 (15)	5,256 (15)	6,986 (15)
Testosterone - pmol/L	12.7 (10.5-15.2)	11.4 (9.4-13.7)	10.4 (8.5-12.6)	0.9 (0.7-1.3)	1.0 (0.7-1.3)	1.0 (0.7-1.4)
Testosterone missing value - yes, n (%)	4,882 (6)	2,897 (6)	2,895 (6)	13,080 (24)	8,261 (23)	10,116 (22)
Missing for any of the biomarkers - yes, n (%)	15,223 (18)	9,074 (18)	8,130 (18)	18,919 (35)	12,077 (34)	15,121 (33)

CSE Certificate of Secondary Education; GCSE General Certificate of Secondary Education; NVQ National Vocational Education; HND Higher National Diploma; HNC Higher National Certificate; MET Metabolic Equivalent of Task; W women; M men

Adiposity-biomarker associations: For men with waist circumference >102 *vs* ≤94 cm, an inverse association was observed for SHBG (GMR 0.79; 95%CI, 0.79-0.79) and testosterone (GMR 0.83; 95%CI, 0.82-0.83); a positive association for CRP (GMR 2.00; 95%CI, 1.97-2.02); and a weak positive association for HbA1c (GMR 1.04; 95%CI, 1.04-1.05). For women with waist circumference >88 *vs* ≤80 cm, an inverse association was observed for SHBG (GMR 0.63; 95%CI, 0.63-0.63); a positive association for CRP (GMR 2.64; 95%CI, 2.60-2.67); and weak positive associations for testosterone (GMR 1.08; 95%CI, 1.07-1.09) and HbA1c (GMR 1.04; 95%CI, 1.04-1.04) (Table 5-2). The associations were similar for BMI and waist-hip ratio in both men and women. (See Supplementary Table A-4 for complete case analysis).

Table 5-2 Association between adiposity measures and biomarkers; multiple imputation analyses excluding women in the middle category of each adiposity measure

		Ratio of geometric means (95% confidence interval)								
		Waist circumference >102 vs. ≤94 M >88 vs. ≤80 F			**Body Mass Index** ≥30 vs. <25 kg/m2			**Waist-Hip Ratio** >0.98 vs. ≤89 M >0.86 vs. ≤77 F		
	Biomarker									
Men		N=128,708			N=88,756			N=88,626		
	CRP	2.00	(1.97 to	2.02)	2.31	(2.28 to	2.34)	2.10	(2.07 to	2.13)
	HbA1C	1.04	(1.04 to	1.05)	1.05	(1.05 to	1.05)	1.05	(1.05 to	1.05)
	SHBG	0.79	(0.79 to	0.79)	0.72	(0.72 to	0.73)	0.77	(0.76 to	0.77)
	Testosterone	0.83	(0.82 to	0.83)	0.80	(0.80 to	0.80)	0.83	(0.83 to	0.84)
Women		N=100,758			N=82,957			N=66,082		
	CRP	2.64	(2.60 to	2.67)	3.19	(3.15 to	3.24)	2.19	(2.16 to	2.23)
	HbA1C	1.04	(1.04 to	1.04)	1.04	(1.04 to	1.04)	1.05	(1.05 to	1.05)
	SHBG	0.63	(0.63 to	0.63)	0.61	(0.60 to	0.61)	0.63	(0.62 to	0.63)
	Testosterone	1.08	(1.07 to	1.09)	1.12	(1.11 to	1.13)	1.04	(1.03 to	1.04)

Abbreviations: CRP C-reactive protein; HbA1c glycated hemoglobin; SHBG sex hormone binding globulin
Missing biomarker data were multiply imputed. Models were adjusted for age at recruitment, education, Townsend deprivation index at recruitment, first-degree relative history of colorectal cancer, physical activity, alcohol intake, smoking status, red and processed meat intake, and regular use of aspirin or ibuprofen. Models for women were additionally adjusted for the number of live births, ever use of oral contraceptive pill, and ever use of hormone therapy.

Adiposity-CRC & Biomarker-CRC associations: The adjusted OR for CRC risk for men with waist circumference >102 *vs* ≤94 cm was 1.38 (95% CI, 1.20-1.59). Similar associations were observed for BMI and waist-hip ratio. A positive association was observed for CRP after adjusting for confounders and waist circumference (OR per doubling concentration 1.07; 95% CI, 1.02-1.11), and there was suggestion for an inverse

association for SHBG (OR 0.92; 95% CI, 0.82-1.03) and testosterone (OR 0.92; 95% CI, 0.81-1.05) after additionally adjusting for CRP, and HbA1c. There was no evidence for an association between HbA1c and CRC risk (OR additionally adjusted for CRP 0.82; 95% CI, 0.58-1.14) (Table 5-3 multiple imputation analysis; Supplementary Table A-5 complete case analysis). No strong evidence for departure from linearity for adiposity-CRC or biomarker-CRC associations was observed (Table 5-3).

For women with waist circumference >88 *vs* ≤80 cm, the adjusted OR for CRC was 1.28 (95% CI, 1.08-1.51). A weaker association was observed for BMI (30 *vs* <25 kg/m^2, OR 1.11; 95% CI, 0.91-1.36) and a stronger association for waist-hip ratio (>0.89 *vs* ≤0.77, OR 1.33; 95% CI, 1.07-1.65). The point estimates and 95% CI were suggestive of weak-to-null associations between biomarkers and CRC (adjusted OR per doubling CRP 1.03; 95%CI, 0.98-1.09; HbA1C 1.06; 95% CI, 0.66-1.71; SHBG 1.03, 95%CI, 0.90-1.18; testosterone 1.00; 95%CI, 0.88-1.13) (Table 5-3 multiple imputation analysis; Supplementary Table A-5 complete case analysis). There was evidence for a non-linear SHBG-CRC association for women. Therefore, in the models for mediation analysis for women, SHBG was included as restricted cubic splines.

Table 5-3 Estimated odds ratios and 95% confidence intervals for the association between adiposity measures (exposure) or biomarkers (mediators) and colorectal cancer (outcome) in men and women; multiple imputation analysis

	Men				Women			
Waist Circumference, cm	**>102 vs <94**				**>88 vs <80**			
OR (95% CI)	1.38	(1.20 to	1.59)		1.28	(1.08 to	1.51)	
Body Mass Index, kg/m^2	**≥30 vs <25**				**≥30 vs <25**			
OR (95% CI)	1.37	(1.16 to	1.62)		1.11	(0.91 to	1.36)	
Waist-Hip ratio	**>0.98 vs <0.89**				**>0.86 vs <0.77**			
OR (95% CI)	1.42	(1.19 to	1.70)		1.33	(1.07 to	1.65)	
Biomarker	**Per doubling concentration**			**P value for evidence against linearity**	**Per doubling concentration**			**P value for evidence against linearity**
CRP								
Model 1	1.07	(1.02 to	1.11)	0.02	1.03	(0.98 to	1.09)	0.75
HbA1C								
Model 1	0.85	(0.61 to	1.19)		1.08	(0.67 to	1.75)	
Model 2 (additionally adjusted for CRP)	0.82	(0.58 to	1.14)	0.15	1.06	(0.66 to	1.71)	0.75
SHBG								
Model 1	0.92	(0.82 to	1.02)		1.01	(0.89 to	1.15)	
Model 2 (additionally adjusted for CRP and HbA1c)	0.92	(0.82 to	1.03)	0.20	1.03	(0.90 to	1.18)	0.00
Testosterone								
Model 1	0.91	(0.80 to	1.04)		1.00	(0.88 to	1.13)	
Model 2 (additionally adjusted for CRP and HbA1c)	0.92	(0.81 to	1.05)	0.56	1.00	(0.88 to	1.13)	0.19

Abbreviations: CRP C-reactive protein; HbA1c glycated hemoglobin; SHBG sex hormone binding globulin
Missing biomarker data were multiply imputed. Models were adjusted for age at recruitment, education, Townsend deprivation index at recruitment, first-degree relative history of colorectal cancer, physical activity, alcohol intake, smoking status, red and processed meat intake, and regular use of aspirin or ibuprofen. Models for women were additionally adjusted for the number of live births, ever use of oral contraceptive pill, and ever use of hormone therapy. Model 1 for biomarkers were additionally adjusted for waist circumference.

Sequential mediation analysis: For men with waist circumference >102 *vs* ≤94 cm, the RR^{NIE} was 1.08 (95% CI, 1.01-1.16) through all the biomarkers, 1.06 (95% CI, 1.01-1.11) through CRP and pathways influenced by CRP, 0.99 (95% CI, 0.97-1.01) through HbA1c excluding the influence of CRP on HbA1c levels, and 1.03 (95% CI, 0.99-1.08) through SHBG and testosterone combined excluding the influences of CRP and HbA1c on SHBG and testosterone levels. The RR^{NDE} not through any of the included biomarkers was 1.26 (95% CI, 1.09-1.47). Similar patterns of mediation were observed for BMI and waist-hip ratio (Table 5-4, see Supplementary Table A-6 for complete case analysis).

For women, the RR^{NIE} was 1.08 (95% CI, 0.95-1.22) through all the biomarkers, 1.08 (95% CI, 0.99-1.17) through pathways influenced by CRP, 1.00 (95% CI, 0.98-1.02) through HbA1c excluding the influence of CRP, and 1.00 (95% CI, 0.92-1.09) through SHBG and testosterone combined excluding the influences of CRP and HbA1c. The RR^{NDE} not through any of the included biomarkers was 1.18 (95% CI, 0.96-1.45). Similar mediation effects were observed for waist-hip ratio. We did not attempt to perform mediation analysis for BMI, because no total effect of BMI on CRC was observed in the multiple imputation (Table 5-4) or complete case analyses (Supplementary Table A-6).

Table 5-4 Estimated natural direct and indirect effects using sequential mediation analysis for the association between adiposity measures and colorectal cancer for men and women, analyses exclude participants categorized as overweight and do not include any interaction terms between adiposity measures and biomarkers (exposure-mediator interactions) or biomarkers (mediator-mediator interactions) - multiple imputation analysis

		Risk Ratio (95% CI)					
	Effect	**Men**			**Women**		
Waist Circumference >102 vs. ≤94 cm M >88 vs. ≤80 cm F	No. Colorectal Cancer affected/ No. unaffected	878/127,830			616/100,142		
	Total effect	1.37	(1.19 to	1.58)	1.27	(1.07 to	1.50)
	Natural indirect effect through all the mediators	1.08	(1.01 to	1.16)	1.08	(0.95 to	1.22)
	Natural indirect effect through CRP	1.06	(1.01 to	1.11)	1.08	(0.99 to	1.17)
	Natural indirect effect through HbA1C, excluding the influence of CRP	0.99	(0.97 to	1.01)	1.00	(0.98 to	1.02)
	Natural indirect effect through SHBG and testosterone, excluding the influence of CRP and HbA1c	1.03	(0.99 to	1.08)	1.00	(0.92 to	1.09)
	Natural direct effect not through any of the mediators	1.26	(1.09 to	1.47)	1.18	(0.96 to	1.45)
Body Mass Index ≥30 vs. <25 kg/m2	No. Colorectal Cancer affected/ No. unaffected	614/88,142			451/82,506		
	Total effect	1.36	(1.15 to	1.60)	1.12	(0.92 to	1.38)
	Natural indirect effect through all the mediators	1.11	(1.00 to	1.22)	-	-	-
	Natural indirect effect through CRP	1.06	(0.99 to	1.13)	-	-	-
	Natural indirect effect through HbA1C, excluding the influence of CRP	0.99	(0.96 to	1.02)	-	-	-
	Natural indirect effect through SHBG and testosterone, excluding the influence of CRP and HbA1c	1.05	(0.99 to	1.12)	-	-	-
	Natural direct effect not through any of the mediators	1.23	(1.02 to	1.48)	-	-	-
Waist-Hip Ratio >0.98 vs. <0.89 M >0.86 vs. ≤0.77 F	No. Colorectal Cancer affected/ No. unaffected	600/88,026			378/65,704		
	Total effect	1.31	(1.09 to	1.58)	1.33	(1.06 to	1.66)
	Natural indirect effect through all the mediators	1.12	(1.03 to	1.23)	1.10	(0.96 to	1.26)
	Natural indirect effect through CRP	1.07	(1.00 to	1.14)	1.09	(1.00 to	1.19)
	Natural indirect effect through HbA1C, excluding the influence of CRP	0.99	(0.96 to	1.02)	1.01	(0.97 to	1.05)
	Natural indirect effect through SHBG and testosterone, excluding the influence of CRP and HbA1c	1.06	(1.01 to	1.11)	1.00	(0.91 to	1.11)
	Natural direct effect not through any of the mediators	1.17	(0.95 to	1.44)	1.20	(0.92 to	1.57)

Abbreviations: CRP C-reactive protein; HbA1c glycated hemoglobin; SHBG sex hormone binding globulin
Models were adjusted for age at recruitment, education, Townsend deprivation index at recruitment, first-degree relative history of colorectal cancer, physical activity, alcohol intake, smoking status, red and processed meat intake, and regular use of aspirin or ibuprofen. Models for women were additionally adjusted for the number of live births, ever use of oral contraceptives, and ever use of hormone therapy.

Additional analyses

Patterns of mediation were similar in analyses that included all possible adiposity-mediator and mediator-mediator interaction terms (Supplementary Table A-7), when calculated free testosterone was used as the biomarker of sex-steroid hormone pathway (Supplementary Table A-8), in analyses limited to colon cancer cases (Supplementary Table A-9), and in analyses limited to colorectal cancer cases diagnosed after two years of follow-up and non-cases with more than two years of follow-up (Supplementary Table A-10).

Potential influence of missing data

We had missing biomarker data for 18% of men and 35% of women. Baseline characteristics were comparable for men with and without missing biomarker data. A slightly larger proportion of women with missing biomarker data had waist circumference ≤80 cm (41% *vs* 39%) or BMI<25 kg/m^2 (40% *vs* 37%) compared with women with complete biomarker data (Supplementary Table A-11).

For all analyses, results from complete case analyses are presented (Supplementary Table A-4 to Supplementary Table A-10). Multiple imputation and complete case analyses yielded similar results but the 95%CIs from multiple imputation analyses were slightly narrower.

Discussion

For men and postmenopausal women, the effect of adiposity on CRC not explained by the included biomarkers was larger than the estimated indirect effect through all biomarkers. Pathways originating with CRP explained a small proportion of the adiposity-CRC association. In men, a small part of the adiposity-CRC association was explained by adiposity-induced reduced levels of SHBG and testosterone.

To our knowledge, this is one of the first and largest studies that has attempted to quantify the mediating effects of multiple pathways on the adiposity-CRC association, while allowing for correlation between mediators and interrelation between pathways (129). The causal mediation analysis approach was able to handle non-linear mediator-outcome associations and exposure-mediator or mediator-mediator interactions (129). These are important advantages over other statistical methods that have traditionally been used to assess mediation (155). The path-specific indirect effects were estimated using a sequential approach. An unverifiable assumption required by this approach is that the prespecified causal ordering of the pathways is true (129). For this analysis, based on current evidence, we assumed that inflammatory and insulin signalling pathways preceded and possibly influenced SHBG and testosterone levels, while inflammatory pathway preceded dysregulated insulin signalling and glucose homeostasis (8, 149).

We made use of all the relevant biomarkers in the UK Biobank, but measurements were available for only a limited number of biomarkers, which was perhaps the most important limitation of this study. The observed mediating role for the pathways, or lack thereof, would have depended on how well the available biomarkers captured the biological characteristics of each pathway. For all participants, we used measurements for the biomarkers at one time point only. The differences in temporal variability and measurement quality of the biomarkers may have introduced varying degrees of information bias into our estimated associations and influenced the strength of the observed indirect effects. Overall, however, the biomarker measurements were of high quality (all CVs <9%) (142, 143, 152) and, uniquely, were available in nearly all cohort participants. Additionally, for all biomarkers, the ICCs calculated for the subgroup of participants who had measurements available for two time points (approximately four years apart) were >0.55, suggesting that single measurements provided a reasonable

estimate of longer-term exposures. In the UK Biobank, height, weight, waist and hip circumference were measured following a published protocol to reduce measurement error (151). In addition to BMI, we used waist circumference and waist-hip ratio as measures of adiposity, which may perform better in identifying individuals with obesity (13). Our study did not provide evidence for an association between BMI and colorectal cancer risk in women, which was similar to a previous analysis of the UK Biobank data (hazard ratio (HR) for highest vs lowest quintile 1.11; 95%CI, 0.89-1.38)(159), but inconsistent with the results of a recent meta-analysis of studies investigating the association between BMI and colorectal cancer (summary RR for women per 5 kg/m^2 1.05; 95%CI, 1.02-1.08)(34) . It is unclear why we found no association between BMI and colorectal cancer for women. One possibility is that a longer duration of follow-up and a larger number of cases may be required to detect the expected weaker (compared with men) BMI and colorectal cancer association. In support of this supposition, in the European Prospective Investigation into Cancer and Nutrition (EPIC) study after an average of 6.1 years of follow-up (n=1570 cases), no association was found between BMI and colorectal cancer for women(160). However, in a more recent EPIC study that included data after an average of 14.9 years of follow-up (n=6291 cases), the expected positive association between BMI and colorectal cancer for women was observed (161)."

The sequential mediation analysis relied on assumptions of no unmeasured confounding of exposure-outcome, mediator-outcome, exposure-mediator, and mediator-mediator relations (129). We had data on several potential confounders and took those into account in our analyses, but due to the observational nature of this study, residual confounding due to unmeasured or mismeasured confounders cannot be ruled out. For the adiposity-biomarker associations, this study was cross-sectional and, thus, prone to reverse

causation. We excluded CRC cases diagnosed within the first year of follow-up to reduce the possibility of outcome having influenced adiposity or biomarker measurements.

Using causal mediation analysis in a nested case-control study within the Health Professionals Follow-up Study (HPFS), it was observed that a large proportion of the BMI-CRC association (OR per 3.6 kg/m^2 1.40; 95%CI, 1.14-1.73) was explained by all inflammatory (CRP, interleukin-6, tumour necrosis factor receptor 2, and macrophage inhibitory cytokine-1) and metabolic (adiponectin, C-peptide, and soluble leptin receptor (sOB-R)) biomarkers (OR^{NIE} 1.26; 95%CI, 0.97-1.52) (133). The authors repeated the mediation analysis separately for the two biomarker groups, assuming that the pathways were independent. Although the HPFS included more inflammatory biomarkers compared with ours, similar to our results, the observed indirect effect through inflammation was small (OR^{NIE} 1.05; 95%CI, 0.96-1.14) (133). Also, in a nested case-control study within the European Investigation into Cancer and Nutrition, adjusting for CRP only changed the RR for waist circumference-CRC association from 1.33 (95%CI, 1.26-1.41) per 10 cm increment to 1.31 (95%CI, 1.24-1.39) in men and from 1.20 (95%CI, 1.14-1.25) to 1.19 (95%CI, 1.14-1.25) in women (117). Additional studies with a wider range of inflammatory biomarkers are needed to further explore the mediating role of inflammation in adiposity-CRC association.

The HPFS nested case-control study found a larger OR^{NIE} for the metabolic biomarkers (OR^{NIE} 1.24; 95%CI, 0.92-1.55), which was mainly driven by adiponectin and sOB-R (133). When the causal mediation analysis was repeated for individual biomarkers, weak to no mediating effect was observed for C-peptide (a biomarker for insulin signalling; OR^{NIE} 1.05; 95%CI, 0.91-1.18) (133). In a cohort study with 4,032 colon and 2,430 rectal cancer cases, the total BMI-colon (HR per 5 kg/m^2 1.14; 95%CI, 1.10-1.19) and -rectal

cancer (HR 1.09; 95%CI, 1.03-1.15) associations were weakly mediated by triglyceride-glucose index (TyG; a proxy for insulin resistance) (HR^{NIE} 1.03; 95%CI, 1.01-1.04 for colon and HR^{NIE} 1.03; 95%CI, 1.01-1.05 for rectal cancer) (134). The strength of the mediating effect was similar in men and women, and after excluding participants who did not have diabetes (134). In our study, we did not observe any mediating effect through HbA1c. The observed indirect effect through TyG in the study by Fritz *et al*. may have been overestimated because the analysis did not allow for potential confounding by the preceding inflammatory pathway (134). In our study, the estimated indirect effect through HbA1c excluded the possible influence of CRP on HbA1c, thus removed the confounding effect of CRP (129). This issue, however, is unlikely to fully explain the null indirect effect for HbA1c we observed. In our study of men and postmenopausal women who at baseline did not have diabetes and did not report taking diabetes medication, there was no evidence for an association between HbA1c and CRC, regardless of adjustment for CRP. Similarly, a recent Mendelian randomisation analysis found no evidence of an effect between genetically-predicted HbA1c concentration and colorectal cancer risk (OR per genetically determined 1 standard deviation (SD) increment 1.02; 95%CI, 0.85 – 1.22) (35). Fritz *et al.* observed an association between TyG and colon (HR per 1 SD 1.07; 95% CI, 1.03-1.10) and rectal (HR 1.09; 95% CI, 1.04-1.14) cancers (134). Although HbA1c is an established long-term biomarker of glucose and insulin metabolism (162), it may not be the ideal biomarker of dysregulated insulin signalling on its own.

We are not aware of other studies that have investigated the mediating role of sex-steroid hormones in the adiposity-CRC association. We were also limited in exploring the role of this pathway fully because we did not have measures for estrogens. Our results provided evidence for a small role of adiposity-induced reduced levels of SHBG and testosterone in explaining the increased CRC risk in men with obesity. This observation

is in line with the existing evidence on the potential protective role of androgens on CRC development (18). It may also partly explain the stronger adiposity-CRC association observed in men because as found in ours and previous studies, adiposity is associated with reduced testosterone levels in men, but not in women (18).

In summary, increased CRP and reduced SHBG and testosterone levels explained a small proportion of the effect of adiposity on CRC in men. Increased CRP also had a small mediating effect on the association in postmenopausal women, comparable to the mediating effect observed in men. We did not observe any mediating effect for HbA1c. A large proportion of the effect of adiposity on CRC was not explained by the included biomarkers. These results suggest that pathways other than those captured here may also be important for understanding the mechanisms linking adiposity and CRC and for identifying targets for CRC prevention in individuals with obesity. Future studies with a wider range of biomarkers reflecting inflammatory status, impaired insulin signalling and hyperinsulinemia, and altered sex-steroid hormone levels in the obese state are needed to perform a comprehensive evaluation of the role of these pathways in explaining the effect of adiposity on CRC link. Proper causal mediation analysis approaches would allow such investigation, while appropriately accounting for correlations between biomarkers.

Acknowledgements: This research was conducted using the UK Biobank Resource under application number 25897.

Disclaimer: Where authors are identified as personnel of the International Agency for Research on Cancer / World Health Organization, the authors alone are responsible for the views expressed in this article and they do not necessarily represent the decisions, policy or views of the International Agency for Research on Cancer / World Health Organization.

Data availability: UK Biobank is an open-access resource. Bonafide researchers can apply to use the UK Biobank dataset by registering and applying at http://ukbiobank.ac.uk/register-apply/.

Financial Support: This work was supported by Australian National Health and Medical Research Council grants (1150591 to SGD and 1104975 to JAS); Melbourne Research Scholarship (to SGD), and the French National Cancer Institute (INCa, grant nr. 2017-127 to NM). MMB is the recipient of an Australian Research Council Discovery Early Career Award (project number DE190101326) funded by the Australian Government. The work reported in this paper was undertaken by SGD while hosted by the International Agency for Research on Cancer (IARC) and funded partially by a PhD Scholarship by IARC. The funders did not influence the design of the study, analysis or interpretation of the data, the writing of this report or the decision to publish.

Conflict of interest: The authors declare no potential conflicts of interest.

Chapter 6: Adiposity and estrogen receptor-positive breast cancer risk in postmenopausal women in the MCCS

In this chapter, an analysis of a case-cohort study within the Melbourne Collaborative Cohort Study (MCCS) has been carried out to investigate the mediating roles of fasting insulin (to reflect hyperinsulinemia) and calculated free estradiol (to reflect alterations in sex-steroid hormone levels) in the effect of adiposity on estrogen receptor (ER)-positive breast cancer risk in postmenopausal women. This chapter includes the author accepted version of the following publication:

Dashti SG, Simpson JA, Karahalios A, Viallon V, Moreno-Betancur M, Gurrin LC, MacInnis RJ, Lynch BM, Baglietto L, Morris HA, Gunter MJ, Ferrari P, Milne RL, Giles GG, English DR. Adiposity and estrogen receptor-positive, postmenopausal breast cancer risk: Quantification of the mediating effects of fasting insulin and free estradiol. Int J Cancer. 2020;146(6):1541-52. DOI: 10.1002/ijc.32504

For consistency, spelling, referencing, tables, and figure numbers have been modified. Supplementary material related to this publication are presented in Appendix B.

Adiposity and estrogen receptor-positive, postmenopausal breast cancer risk: quantification of the mediating effects of fasting insulin and free estradiol

S Ghazaleh Dashti[1,2,3], Julie A Simpson[1], Amalia Karahalios[1,4], Vivian Viallon[3], Margarita Moreno-Betancur[1,5], Lyle C Gurrin[1], Robert J MacInnis[1,2], Brigid M Lynch[1,2,6], Laura Baglietto[7], Howard A Morris[84], Marc J. Gunter[3], Pietro Ferrari[3], Roger L Milne[1,2], Graham G Giles[1,2], Dallas R English[1,2]

1 Centre for Epidemiology and Biostatistics, Melbourne School of Population and Global Health, The University of Melbourne, Melbourne, Victoria, Australia
2 Cancer Epidemiology and Intelligence Division, Cancer Council Victoria, Melbourne, Victoria, Australia
3 Section of Nutrition and Metabolism, International Agency for Research on Cancer (IARC), Lyon, France
4 School of Public Health and Preventive Medicine, Monash University, Melbourne, Victoria, Australia
5 Clinical Epidemiology and Biostatistics Unit, Murdoch Children's Research Institute, Melbourne, Australia
6 Physical Activity Laboratory, Baker Heart and Diabetes Institute, Melbourne, Victoria, Australia
7 Department of Clinical and Experimental Medicine, University of Pisa, Italy
8 School of Pharmacy and Medical Sciences, University of South Australia, Adelaide, Australia

Corresponding author
Dallas English
Centre for Epidemiology and Biostatistics
Melbourne School of Population and Global Health
Level 3, 207 Bouverie Street
The University of Melbourne VIC 3010 Australia
Phone: +61 3 8344 0742
Email: d.english@unimelb.edu.au

Short title: *Adiposity and breast cancer: a causal mediation* analysis

Key words: adiposity, estrogen receptor positive, postmenopausal breast cancer, causal mediation analysis, insulin, estrogens, free estradiol, circulating estrogens

Article category: Cancer Epidemiology

[4] Deceased 18 April 2019

Abbreviations used: ER: estrogen receptor; SHBG: sex hormone binding globulin; BMI: body mass index; RR: relative risk
CI: confidence interval; IGF-1: insulin-like growth factor-1; WHI-OS: Women's Health Initiative Observational Study; HR: hazard ratio; PLCO: Prostate, Lung, Colorectal, and Ovarian Cancer; MCCS: Melbourne Collaborative Cohort Study; ICC: intraclass correlation; TE: total effect; IDE: interventional direct effect; IIE: interventional indirect effect; IQR: interquartile range

Novelty and Impact: Adiposity increases the risk of estrogen receptor-positive postmenopausal breast cancer. Better understanding of mechanisms underlying this association could help with identifying new targets for cancer prevention. In this study of 1,178 women (149 cases), we applied novel statistical methods to quantify mediating effects of insulin and free estradiol on this association. Our results suggest that free estradiol explains adiposity-breast cancer association only in part. Identifying other pathways contributing to this relation should be a priority.

Abstract

Adiposity increases estrogen receptor (ER)-positive postmenopausal breast cancer risk. While mechanisms underlying this relation are uncertain, dysregulated sex-steroid hormone production and insulin signalling are likely pathways. Our aim was to quantify mediating effects of fasting insulin and free estradiol in the adiposity and ER-positive postmenopausal breast cancer association. We used data from a case-cohort study of sex hormones and insulin signalling nested within the Melbourne Collaborative Cohort Study. Eligible women, at baseline, were not diagnosed with cancer, were postmenopausal, did not use hormone therapy, and had no history of diabetes or diabetes medication use. Women with ER-negative disease or breast cancer diagnosis within the first follow-up year were excluded. We analysed the study as a cumulative sampling case-control study with 149 cases and 1,029 controls. Missing values for insulin and free estradiol were multiply imputed with chained equations. Interventional direct (IDE) and indirect (IIE) effects were estimated using regression-based multiple-mediator approach. For women with body mass index (BMI) >30kg/m^2 compared with women with BMI 18.5kg/m^2 to 25kg/m^2, the risk ratio (RR) of breast cancer was 1.75 (95% confidence interval (CI) 1.05-2.91). The estimated IDE (RR) not through the mediators was 1.03 (95% CI 0.43-2.48). Percentage mediated effect through free estradiol was 72% (IIE-RR 1.56; 95% CI 1.11-2.19). There was no evidence for an indirect effect through insulin (IIE-RR 1.12; 95% CI 0.68-1.84; 28% percentage mediated). Our results suggest that circulating free estradiol plays an important mediating role in the adiposity-breast cancer relationship but does not explain all of the association.

Introduction

There is convincing evidence that adiposity increases the risk of estrogen receptor (ER)-positive breast cancer in postmenopausal women (30, 41). The mechanisms underlying this association are uncertain, but disturbed production of sex-steroid hormones related to increased mass and activity of adipose tissue is recognized as one of the main candidate pathways (163). Circulating estrogen concentrations are elevated in overweight and obese postmenopausal women while sex hormone binding globulin (SHBG) concentrations are decreased, thereby raising bioavailable estrogens (estradiol and estrone) (13, 164). Estrogens might increase breast cancer risk through their oxidative metabolites, and by affecting cell proliferation and apoptosis through their interaction with the ER in breast tissue (165). A positive association between circulating estrogens and the risk of postmenopausal breast cancer has been documented consistently by prospective epidemiological studies (111, 138, 166). In addition, suggestive of a mediating effect of estrogens on the adiposity-postmenopausal breast cancer association, the association between body mass index (BMI) and postmenopausal breast cancer is attenuated after adjusting for circulating estrogens, in particular, free estradiol (121, 167). A pooled analysis of eight prospective studies showed that adjusting for free estradiol changed the relative risk (RR) of postmenopausal breast cancer for a 5 kg/m^2 increment in BMI from 1.19 (95% confidence interval (CI) 1.05 to 1.34) to 1.02 (95% CI 0.89 to 1.17) in women not using hormone therapy at the time of blood draw (167).

Insulin resistance, a condition with a well-established association with adiposity, is another pathway hypothesized to link adiposity with postmenopausal breast cancer (163). Insulin could affect cancer risk directly by its mitogenic properties and by increasing secretion of insulin-like growth factor-1 (IGF-1) and down-regulating IGF-1 binding proteins (14). Some evidence suggests that insulin might increase breast cancer risk by

downregulating SHBG production and increasing bioavailable estrogen levels (12, 163). Epidemiological studies investigating the association between insulin or C-peptide (a marker for insulin production) and postmenopausal breast cancer risk have been inconclusive (91, 163). However, in a case-cohort study within the Women's Health Initiative Observational Study (WHI-OS), a positive association was observed between fasting insulin and postmenopausal breast cancer in women not using hormone therapy and adjusting for fasting insulin substantially attenuated the association between BMI and postmenopausal breast cancer risk (hazard ratio (HR) for BMI≥30 kg/m^2 compared with 18.5 to 24.9 kg/m^2 prior to and after adjustment were 2.12 95% CI 1.26 to 3.58) and 1.50 (95% CI 0.80 to 2.83) respectively) (119).

Quantification of the effects of biomarkers as intermediaries of the adiposity-postmenopausal breast cancer association using formal mediation analysis might contribute to a better understanding of the potential mechanistic pathways (155). This knowledge might be instrumental in identifying targets for pharmacological and lifestyle interventions for the prevention of adiposity-related breast cancer (13). Two studies within the WHI-OS (120) and the Prostate, Lung, Colorectal, and Ovarian Cancer (PLCO) study (135) have used modern methods for mediation analysis based on the potential outcomes approach (155) to quantify the mediating effects of insulin (120) and estradiol (120, 135) in the BMI-breast cancer association. Neither of these studies had a measure of free estradiol, which, compared with total estradiol, might play a more important role in mediating the effect of adiposity on breast cancer risk (121, 167).

For this study, we used data from a case-cohort study of sex hormones and insulin signalling and cancer nested within the Melbourne Collaborative Cohort Study (MCCS) (138, 168). We used a regression-based multiple-mediator approach to estimate

interventional indirect effects (130) quantifying the extent to which the associations between two proxy measures for adiposity - BMI and waist circumference – and ER-positive postmenopausal breast cancer were mediated through fasting insulin and free estradiol.

Materials and Methods

The Melbourne Collaborative Cohort Study

The MCCS includes 24,469 women recruited in Melbourne, Australia, in 1990-1994 (137). Almost all participants (99%) were aged between 40 and 69 years at baseline. Clinical data collection, including blood sample (74% fasting) and height, weight, and waist circumference measurements were administered at baseline based on standard procedures by nurses at dedicated recruitment clinics (137). Plasma was stored in liquid nitrogen. Women were interviewed about lifestyle factors, use of hormone therapy and oral contraceptives, and reproductive history. The study protocol was approved by the Cancer Council Victoria Human Research Ethics Committee and participants provided written informed consent to participate. The data that support the findings of this study are held by the Cancer Epidemiology Division, Cancer Council Victoria. Restrictions apply to the availability of these data, which were used under license for this study. Information on how data from the MCCS (Health 2020) data can be accessed is available at https://www.pedigree.org.au/.

Participants

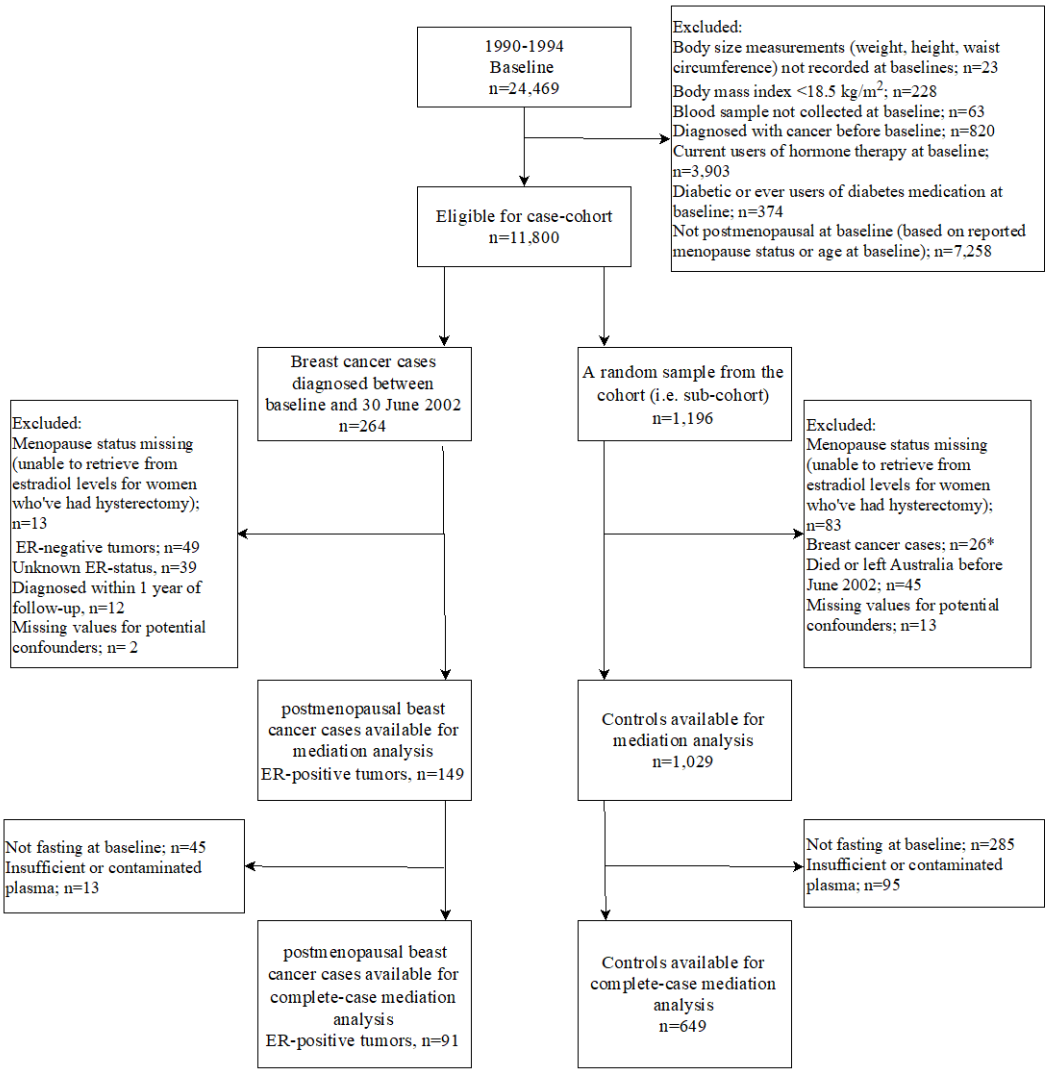
Details of the case-cohort study have been published (137). For this analysis, women were ineligible if, at baseline, they did not have body size measurements recorded or did not provide blood, had BMI<18.5kg/m2, had been diagnosed with any cancer, were current users of hormone therapy, had a history of diabetes or diabetes medication use, or

were not postmenopausal. Women were classified as postmenopausal if their periods had stopped naturally for more than 12 months, or their menopausal status was unknown but they were at least 56 years old (169), or they had a hormone measurement recorded, with estradiol concentration <109 pmol/L (169). The sample included all eligible women diagnosed with invasive adenocarcinoma of the breast between baseline and 30 June 2002 (although for this analysis women with ER-negative or unknown ER status cancers were excluded), and a random sample of eligible women from the cohort (sub-cohort) (Figure 6-1).

Diagnoses of breast cancer were ascertained through record linkage to the population-based Victorian Cancer Registry. ER status was determined from histopathology reports or, if not reported, from immunohistochemical analysis of archival tumour tissue (138).

For the mediation analysis, because methods for handling multiple mediators with survival outcomes or complex designs have only recently been proposed and are still being developed (144), we analysed the study as a case-control study by using as the control group women in the sub-cohort who had not been diagnosed with breast cancer (Figure 6-1). Sub-cohort members who left Australia or died before 30 June 2002 (identified by record linkage to Victorian death records and the National Death Index) were also excluded. Hence, the remaining sub-cohort participants were known to be genuine “controls” and constituted the control group of the case-control analysis.

Figure 6-1 Flowchart of female participants in the Melbourne Collaborative Cohort Study; abbreviation: ER estrogen receptor



For the purpose of mediation analysis women in the sub-cohort who were diagnosed with breast cancer were not included as controls, but they were kept as cases if they met the eligibility criteria.

Plasma Analysis

Free estradiol concentration was calculated from measured circulating total estradiol and SHBG. Insulin, total estradiol, and SHBG were not measured for women who had insufficient or contaminated plasma (Figure 6-1). Insulin was only measured for women who were fasting at the time of blood draw (Figure 6-1). A detailed description of the laboratory measurements has been published (138). Briefly, plasma samples were retrieved from liquid nitrogen, aliquoted into 450-µL amounts, randomly allocated to batches with approximately equal number of cases and sub-cohort participants across

batches, and shipped to the laboratory of H.M. One scientist, blinded to the case-control status of the samples, performed all the measurements. Samples were thawed in a warm water bath, vortexed rapidly for a few seconds, and centrifuged at 2,000 (210 × g) for 10 minutes. The assays used for measurement of total estradiol and SHBG were electrochemiluminescence immunoassay (Elecsys 2010 analyzer, Roche Diagnostics GmbH, Mannheim, Germany) and immunometric assay (IMMULITE analyzer, DPC, Los Angeles, CA, USA) respectively. The overall coefficients of variation from pooled plasma were 10% at 157 pmol/L for total estradiol and 7% at 45.0 nmol/L for SHBG. From a reliability study done in 45 women with blood samples drawn at baseline and approximately 1 year later, the intraclass correlation (ICC) was 0.90 (0.85 – 0.95) for SHBG; the ICC was not estimated for total estradiol because of insufficient samples (138). Free (protein-unbound) estradiol concentration was calculated from measured total estradiol and SHBG, assuming a fixed albumin concentration of 40 g/L and using the law of mass action (138).

For the fasting participants, plasma insulin was measured using AxSYM Microparticle Enzyme Immunoassay (Abbott, North Ryde, NSW, Australia).

Statistical Analysis

Waist circumference and BMI were used as proxy measures for adiposity in separate analyses. The World Health Organisation cut-off values were used to generate binary variables comparing obese with normal weight (BMI, ≥30 vs <25 kg/m^2; waist circumference, >88 vs ≤80 cm) for each measure. The choice to primarily treat these exposure variables as binary was made to assist with the interpretation of the estimated mediated effects (see Mediation Analysis) and utility for policy making.

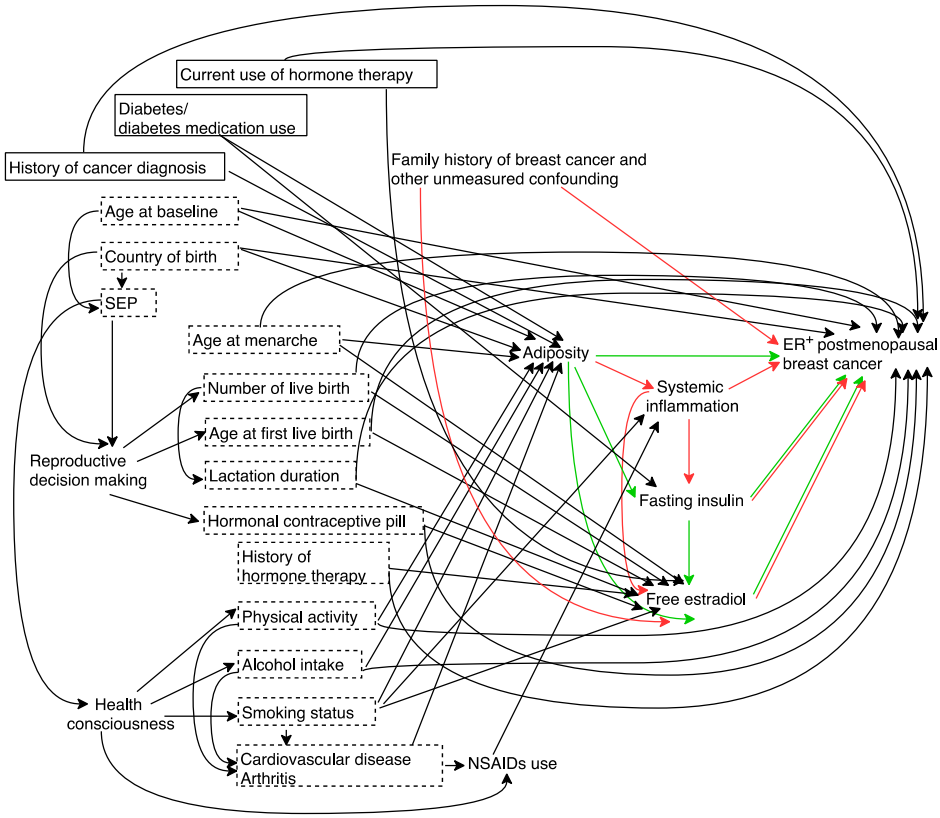
To remove batch effects, we first fitted linear mixed effects models, including a random effect for batch, to the log-transformed biomarker values for women in the sub-cohort.

Next, for all women, the predicted batch-specific deviations from the overall mean were subtracted from the observed values.

Exposure-mediator association: Linear regression models were fitted to describe the relationship between fasting insulin and calculated free estradiol concentrations and adiposity quantified as BMI or waist circumference; for each biomarker, ratios of geometric means and corresponding 95% CIs were calculated for women classified as overweight/obese compared with those classified as normal BMI and waist circumference. Models were fitted to the sub-cohort participants with available biomarker measurements (complete case) and adjusted for potential confounders identified using prior knowledge in the literature and depicted in a causal diagram (Figure 6-2).

Exposure-outcome and mediator-outcome associations: Cox regression with age as the time axis was fitted to estimate the HR and 95% CI for the association between postmenopausal breast cancer and each adiposity measure or biomarker. The Prentice approach with robust standard errors was used to take the case-cohort design into account (170). For the sub-cohort, observation time started at age at recruitment and ended at breast cancer diagnosis, death, last age known to be in Australia, or the end of follow-up, whichever occurred first. All models included potential confounders (Figure 6-2). Graphical inspection of the Schoenfeld residuals suggested no departure from proportionality of the hazards (171).

Figure 6-2 Assumed causal diagram for quantifying the mediating effects of fasting insulin, and free estradiol on adiposity- estrogen receptor-positive postmenopausal breast cancer association; abbreviations: ER estrogen receptor



To simplify the diagram of the assumed causal relations between adiposity (exposure), estrogen receptor-positive postmenopausal breast cancer (outcome), fasting insulin and free estradiol (mediators) and potential confounders, we only included the arrows that were sufficient to flag a variable as a common cause (i.e. confounder) of exposure-outcome, or exposure-mediator, or mediator-outcome. Therefore, this is not a causal diagram *per se*. It is assumed that all potential confounders preceded the exposure, mediators, and the outcome. The diagram was developed with reference to the literature (41, 172, 173) and expert opinion. The green arrows depict the direct and indirect pathways we were interested in, the red arrows the pathways that might have introduced biased due to unmeasured confounding, and the black arrows the backdoor pathways that were blocked by conditioning on the boxed variables. Dashed boxes indicate variables that were included in multivariable analyses, and solid boxes variables that were used as eligibility criteria. Based on the assumed causal ordering of variables in our diagram, no collider bias would have been introduced by conditioning on any of the identified confounders in the figure.

Mediation Analysis with Multiple Mediators: For mediation analysis, the binary case-control status was the outcome. The exposure was the binary measure of adiposity; mediators were the continuous normalized variables for fasting insulin and calculated free estradiol. All mediation analyses were guided by the causal diagram in Figure 6-2, which was developed with reference to the existing literature (41, 172, 173) as well as expert opinion and took into account the potential confounders of the exposure-outcome, mediator-outcome, and exposure-mediator relationships (155). The total effect (TE) of

adiposity on the risk of ER-positive breast cancer was estimated using logistic regression, with the result in this case-control analysis providing similar results to those from the original case-cohort study using Cox regression. The TE was decomposed into the interventional direct effect (IDE) not through either of the mediators, the interventional indirect effect (IIE) through fasting insulin, IIE through calculated free estradiol, and the difference between the TE and the sum of the estimated IDE and IIE (130). This latter effect essentially captures the IIE through the dependence between fasting insulin and calculated free estradiol (130). Table 6-1 summarizes the interpretations and Supplementary Table B-1 the expressions used for estimating the IDE and IIEs, which are based on (130, 132). The effects were estimated on the risk difference and risk ratio scales, based on the following models: 1) linear regression of fasting insulin conditional on exposure and baseline confounders; 2) linear regression model of calculated free estradiol conditional on the exposure and baseline confounders; 3) linear regression model of calculated free estradiol conditional on the exposure, fasting insulin, an interaction term between exposure and fasting insulin, and baseline confounders, 4) logistic regression of the outcome conditional on the exposure, both mediators, interaction terms between exposure and each mediator, interaction term between the two mediators, and baseline confounders. Models 1, 3 and 4 were combined to estimate the IDE, and models 1, 2 and 4 to estimate the IIEs through fasting insulin or calculated free estradiol. The case-control design was taken into account by fitting the models for the mediators to the sub-cohort (128) and by subtracting the log of (1/sampling fraction of controls) from the constant term of the outcome model prior to calculating risks (174).

Using the estimated absolute risk differences, we further calculated the proportion mediated, which could be interpreted as the importance of the biomarker in explaining

the adiposity-cancer risk, by dividing the IIE risk difference by the sum of IDE and IIE risk differences (19).

Table 6-1 The interpretations of interventional direct and indirect effects estimated in this study

#	Effect	Interpretation [20,26]
1	Interventional direct effect not through fasting insulin or calculated free estradiol	Average relative change in probability of ER-positive postmenopausal breast cancer risk for overweight or obese versus normal weight women, when fixing the joint distribution of fasting insulin and calculated free estradiol to a random draw from their observed joint distribution adjusted for baseline confounders in normal weight women.
2	Interventional indirect effect through fasting insulin	Average relative change in probability of ER-positive postmenopausal breast cancer risk for overweight or obese women, while intervening to shift the distribution of fasting insulin from its observed distribution in overweight or obese women adjusted for baseline confounders, to its observed distribution in normal weight women adjusted for the baseline confounders and fixing the distribution of calculated free estradiol to a random draw from its observed distribution in normal weight women adjusted for the baseline confounders.
3	Interventional indirect effect through calculated free estradiol	Average relative change in probability of ER-positive postmenopausal breast cancer risk for overweight or obese women, while fixing the distribution of fasting insulin to a random draw from its observed distribution in overweight or obese women, adjusted for the baseline confounders and intervening to shift the distribution of calculated free estradiol from its observed distribution in overweight or obese women adjusted for the baseline confounders to its observed distribution in normal weight women adjusted for the baseline confounders
4	Interventional indirect effect through dependence of fasting insulin and calculated free estradiol	Difference between the total effect and the sum of direct and indirect effects (effects 1 to 3). This effect captures the average relative change in probability of ER-positive postmenopausal breast cancer for overweight or obese women through the dependence between fasting insulin and calculated free estradiol. This effect is generally expected to diverge from null in the presence of mediator-mediator and exposure-mediator interactions (130).

The corresponding expressions used for the effects are provided in Supplementary Table B-1

Missing data: We had missing values for fasting insulin and calculated free estradiol (Figure 6-1). For mediation analysis, we used multiple imputation based on chained equations (175) with 40 iterations to impute missing values for each mediator, as well as for the interaction terms between each mediator and the outcome, and for the interaction term between the two mediators (157). The imputation models were stratified on BMI and waist circumference (in separate models) (157), and included all the variables in the final mediation models and additional auxiliary variables (height, hip circumference, weight, and sub-cohort membership). The imputed mediator variables were then used to generate the interaction terms between each mediator and exposure (157).

Simulation-based regression-standardisation was used to estimate mediation effects separately for each of the imputed datasets. This approach entails averaging estimates based on a large number of Monte-Carlo draws (in our case, 600,000) from the mediator distributions, with the distribution of the confounders set to be the same as their observed distribution in the sample, as suggested in (130). Standard errors for the TE, IDE, and IIE (and 95% CI) within each imputed dataset were estimated using 1000 bootstrap samples. The estimates and standard errors derived for each imputed dataset were then pooled using Rubin rules (146) to calculate the multiple imputation estimate and its variance.

All analyses were performed in Stata version 14 (176).

Sensitivity Analysis: Mediation analyses were repeated to compare postmenopausal breast cancer risk for women with BMI≥25 vs <25 kg/m^2 and waist circumference >80 vs ≤80 cm. We also present results for continuous exposure measures (per 5 kg/m^2 increase in BMI or 10 cm increase in waist circumference).

Results

Of all women in the MCCS who met the eligibility criteria, 1,196 were randomly selected for the case-cohort study (Figure 6-1). Eighty-three women were excluded because of missing menopausal status and 13 women due to missing values for potential confounders. For the case-control analysis 45 women who had died or left Australia before the end of follow-up were additionally excluded. During follow-up, 264 women who met the eligibility criteria were diagnosed with breast cancer; 26 of these were also members of the sub-cohort; in the mediation analysis, these 26 women were included as cases only. Of the breast cancer cases, we excluded 13 with missing menopause status, 88 with ER-negative or unknown ER status tumours, 2 with missing values for potential confounders, and 12 who were diagnosed within the first year after recruitment. The final

sample for mediation analysis comprised 149 ER-positive breast cancer cases and 1,029 controls. Of these, 330 women (45 cases) were not fasting at the time of blood draw and 108 (13 cases) had insufficient or contaminated plasma and were therefore not included in the complete-case analyses involving biomarkers, but they were in multiple imputation analyses.

The median age at cancer diagnosis was 68 ((interquartile range (IQR) 64 to 71) years and the median time to cancer diagnosis from baseline was 5.5 (IQR 3.7 to 7.0) years. For controls, the median follow-up time was 9.3 (IQR 8.4 to 10.3) years. Table 6-2 summarizes the baseline characteristics of participants included in the case-control analysis.

Table 6-2 Baseline characteristics of cases and controls

	Breast cancer cases (N=149)	Controls (N=1,029)
Age at baseline, *years*; median (interquartile range)	63 (59 – 66)	61 (57 – 65)
Was fasting at baseline and had plasma available for biomarker measurement; n (%)	91 (61)	649 (63)
Country of birth, *grouped*; n (%)		
Australia/New Zealand/ Northern Europe	119 (80)	758 (74)
Southern Europe	30 (20)	271 (26)
Socioeconomic position, *quintile*; n (%)		
1 (most disadvantaged)	30 (20)	228 (22)
2	31 (21)	244 (24)
3	23 (15)	133 (13)
4	31 (21)	183 (18)
5 (least disadvantaged)	34 (23)	241 (23)
Physical activity score; n (%)		
None	24 (16)	217 (21)
Low	36 (24)	217 (21)
Medium	60 (40)	407 (40)
High	29 (20)	188 (18)
Smoking status; n (%)		
Never smoked	110 (74)	749 (73)
Former smoker	34 (23)	200 (19)
Current smoker	5 (3)	80 (9)
Alcohol consumed per day in the last decade, *grams per day*; n (%)		
0	69 (46)	562 (55)
1-19	66 (44)	382 (37)
≥20	14 (9)	85 (8)
Personal history of cardiovascular disease; n (%)	13 (9)	57 (6)
Personal history of arthritis; n (%)	67 (45)	493 (48)
Age at menarche, *years*; n (%)		
<12	25 (17%)	141 (13.7%)
12	24 (16%)	210 (20.4%)
13	36 (24%)	263 (25.6%)
>13	64 (43%)	415 (40.3%)
Personal history of contraceptive pill use; n (%)	62 (42)	455 (44.2%)
Personal history of hormone therapy use; n (%)	21 (14)	143 (13.9%)
Parity and age at first live birth; n (%)		
Nulliparous	16 (11)	120 (12)
1	11 (7)	70 (7)
>1 & <25 years	60 (40)	421 (41)
>1 & ≥25 years	62 (42)	418 (41)
Lactation duration; *months*; n (%)		
Never	28 (19)	177 (17)
Up to 6	46 (31)	293 (29)
7-12	27 (18)	174 (17)
>12	48 (32)	385 (37)
Body mass index, *kg/m²*; median (interquartile range)	27.5 (24.4-32.3)	26.7 (24.2-30.0)
≥18.5 - <25	46 (31)	346 (34)
≥25 - <30	52 (35)	425 (41)
≥30	51 (34)	258 (25)
Waist circumference, *cm*; median (interquartile range)	82.5 (76-94.5)	80.9 (74.0-89.0)
≤80	61 (41)	492 (48)
>80 - ≤88	32 (22)	265 (26)
>88	56 (38)	272 (26)
Fasting insulin, *pmol/L*, median (interquartile range)	5.3 (3.8-8.8)	5.5 (3.8-7.7)
Calculated free estradiol, *pmol/L*, median (interquartile range)	0.9 (0.6-1.2)	0.8 (0.6-1.0)

Exposure-mediator associations: The geometric mean fasting insulin was 1.73 (95% CI 1.56 to 1.93) times higher for women with BMI ≥30 kg/m^2 compared with women with

BMI <25 kg/m^2, and the geometric mean for calculated free estradiol was 1.49 (95% CI 1.36 to 1.62) times greater. Similar associations were observed for waist circumference (Table 6-3).

Table 6-3 Estimated relative change in geometric mean of fasting insulin and calculated free estradiol levels associated with measures of adiposity in a random sample of the cohort eligible for this study and available for complete-case analysis (sub-cohort members; n=649; 9 with breast cancer)

	Median value	Relative change in geometric mean (95% CI)	
Exposure		Fasting insulin	Calculated free estradiol
Body mass index, kg/m^2			
<25	23.2 kg/m^2	1 (Reference)	1 (Reference)
≥25 - <30	27.1 kg/m^2	1.25 (1.15-1.38)	1.21 (1.12-1.30)
≥30	33.2	1.73 (1.56-1.93)	1.49 (1.36-1.62)
Waist circumference, cm			
≤80	73.5 cm	1 (Reference)	1 (Reference)
>80 - ≤88	84 cm	1.33 (1.21-1.47)	1.19 (1.10-1.29)
>88	94	1.65 (1.50-1.81)	1.37 (1.26-1.48)

Estimates are from linear regression models of log transformed normalized biomarker values on measure of adiposity (obese/overweight versus normal classifications), adjusted for age at baseline, country of birth, socioeconomic index for area, physical activity, smoking status, alcohol intake, history of cardiovascular disease, history of arthritis, age at menarche, history of hormonal contraceptive use, history of hormone therapy use, parity and age at first live birth, and lactation duration

Exposure-outcome and mediator-outcome associations: The relationship between ER-positive postmenopausal breast cancer risk and adiposity measures and biomarkers are reported in Table 6-4. For both BMI and waist circumference, women categorized as obese had an increased HR for breast cancer compared with the reference group (BMI ≥30 vs <25kg/m^2 HR 1.75; 95% CI 1.10 to 2.77; waist circumference >88 vs ≤80cm hazard ratio 1.85; 95% CI 1.22 to 2.81). There was also an increased risk of postmenopausal breast cancer associated with calculated free estradiol (HR 1.56, 95% CI 1.12 to 2.16), but not with fasting insulin (HR 1.00, 95% CI 0.65to 1.53).

Table 6-4 Association between measures of adiposity, fasting insulin, and calculated free estradiol and ER-positive postmenopausal breast cancer risk

	Case-cohort study (n=1,222) *		Case-control sub-study (n=1,178) †
	No cases/ person-years	Hazard Ratio (95% CI)	Odds Ratio (95% CI)
Body mass index			
<25	46/3,274	1 (Reference)	1 (Reference)
≥25 - <30	52/4,201	0.98 (0.63-1.52)	1.03 (0.67 to 1.60)
≥30	51/2,567	1.75 (1.10-2.77)	1.78 (1.12 to 2.83)
Waist circumference;			
≤80	61/4,676	1 (Reference)	1 (Reference)
>80 - ≤88	32/2,585	1.07 (0.67-1.69)	1.08 (0.68-1.73)
>88	56/2,781	1.85 (1.22-2.81)	1.93 (1.26-2.95)
Biomarkers, *per doubling concentration* ‡	91/6,321		
Fasting insulin		1.00 (0.65 to 1.53)	1.06 (0.72 to 1.57)
Calculated free estradiol		1.56 (1.12 to 2.16)	1.54 (1.07 – 2.22)

* For these analyses, age was the timescale, observation time started at age at recruitment and ended at age at breast cancer diagnosis, death, last known age known to be in Australia, or age at the end of follow-up, whichever occurred first. Prentice method and robust standard errors were used to account for the case-cohort design. † For these analyses, controls who left Australia or died before 30 June 2002 were excluded (n=45). ‡ Estimates are from complete case analysis (i.e. excluding women who did not have biomarker measurements) and models that included log-2 transformed biomarkers values after normalisation and were additionally adjusted for body mass index.
All models were adjusted for country of birth, socioeconomic index for area, physical activity, smoking status, alcohol intake, history of cardiovascular disease, history of arthritis, age at menarche, history of hormonal contraceptive use, history of hormone therapy use, and parity and age at first live birth and lactation duration

Mediation Analysis with Multiple Mediators: From mediation analysis with multiple imputation to handle missing insulin and estradiol data, when using BMI as proxy for obesity, the TE risk ratio for obese women compared with normal women was 1.75 (95% CI 1.05 to 2.91). The estimated IDE risk ratio not through either of the mediators was 1.03 (95% CI 0.43 to 2.48). The risk ratio was 1.12 (95% CI 0.68 to 1.84) through fasting insulin and 1.56 (95% CI 1.11 to 2.19) through calculated free estradiol (point estimate on the risk difference scale corresponding to 28% and 72% mediated effect respectively). The estimated IIE risk ratio through the interdependence between fasting insulin and calculated free estradiol (effect #4 Table 6-1) was 0.90 (95% CI 0.63 to 1.29) (Table 6-5).

When using waist circumference as proxy for obesity, the TE risk ratio for overweight or obese versus normal women was 1.96 (95% CI 1.23 to 3.13); the IDE was 1.02 (95% CI 0.47 to 2.22); and the IIEs were 1.23 (95% CI 0.81 to 1.88) through fasting insulin (32% mediated effect), 1.42 (95% CI 1.09 to 1.87) through calculated free estradiol (54%

mediated effect), and 1.00 (95% CI 0.81 to 1.25) through the interdependence between the two biomarkers (Table 6-5).

Sensitivity analyses: Results comparing ER-positive breast cancer risk for women with BMI≥25 vs <25 kg/m^2 or waist circumference >80 vs ≤80 cm, as well as results for continuous exposures are presented in Supplementary Table B-2. For all exposure variables, there was evidence for a mediated effect through free estradiol (with the proportion mediated ranging from 30% for waist circumference per 10 cm and 46% for waist circumference >80 vs ≤80 cm to 51% for BMI per 5 kg/m^2 and 71% for BMI ≥25 vs <25 kg/m^2) but not for fasting insulin.

Table 6-5 Results of multiple-mediator analysis with fasting insulin and free estradiol as the mediators of interest, body mass index (≥30 vs <25 kg/m^2) or waist circumference (>88 vs ≤80 cm) as exposure, and non-estrogen receptor negative postmenopausal breast cancer as outcome.

		Risk ratio (95% confidence interval)		Risk difference, per 10,000 (95% confidence interval)	
	Effect	Complete case analysis*	Multiple imputation†	Complete case analysis*	Multiple imputation†
BMI ≥30 vs <25 kg/m^2	Total effect	1.17 (0.55 to 2.47)	1.75 (1.05 to 2.91)	18.7 (-56.9 to 94.2)	80.6 (11.5 to 149.8)
	Interventional direct effect not through insulin and free estradiol	0.39 (0.08 to 1.92)	1.03 (0.43 to 2.48)	-51.8 (-114.9 to 11.3)	18.1 (-67.4 to 103.7)
	Interventional indirect effect through insulin	1.21 (0.56 to 2.60)	1.12 (0.68 to 1.84)	25.0 (-66.1 to 116.1)	28.0 (-70.6 to 126.6)
	Interventional indirect effect through free estradiol	2.08 (1.14 to 3.82)	1.56 (1.11 to 2.19)	71.5 (0.9 to 142.1)	72.2 (16.7 to 127.7)
	Interventional indirect effect through the interdependence of insulin and free estradiol	0.99 (0.60 to 1.63)	0.90 (0.63 to 1.29)	-10.8 (-64.3 to 42.8)	-18.7 (-81.3 to 43.9)
Waist Circumference >88 vs ≤80 cm	Total effect	1.15 (0.62 to 2.13)	1.96 (1.23 to 3.13)	18.3 (-51.0 to 87.5)	97.4 (29.6 to 165.1)
	Interventional direct effect not through insulin and free estradiol	0.29 (0.08 to 1.02)	1.02 (0.47 to 2.22)	-53.3 (-107.2 to 0.6)	23.1 (-41.7 to 88.0)
	Interventional indirect effect through insulin	1.42 (0.83 to 2.42)	1.23 (0.81 to 1.88)	27.0 (-48.9 to 102.8)	36.0 (-42.7 to 114.8)
	Interventional indirect effect through free estradiol	1.89 (1.23 to 2.91)	1.42 (1.09 to 1.87)	68.8 (7.3 to 130.3)	61.2 (13.8 to 108.7)
	Interventional indirect effect through the interdependence of insulin and free estradiol	1.14 (0.86 to 1.51)	1.01 (0.81 to 1.25)	-4.0 (-40.6 to 32.6)	-6.8 (-40.0 to 26.4)

*Models for mediation analyses were based on 600,000 Monte Carlo draws and 95% confidence intervals were constructed from 1,000 bootstrap draws. † Mediation analyses were performed within each imputed dataset (including 97 cases and 604 controls for BMI ≥30 vs <25 kg/m^2 and 117 cases and 764 controls for waist circumference >88 vs ≤80 cm) and resulting estimates and standard errors (derived from bootstrap draws) were pooled to obtain the multiple imputation estimates and 95% confidence intervals. All models were adjusted for country of birth, socioeconomic index for area, physical activity, smoking status, alcohol intake, history of cardiovascular disease, history of arthritis, age at menarche, history of hormonal contraceptive use, history of hormone therapy use, parity and age at first live birth, and lactation duration

Potential influence of missing data: Biomarker measurements were not available for 37% of the participants. When compared with women who had missing data for the mediation analysis, women with complete data had similar proportions of breast cancer cases (12% vs. 13%), as well as distributions of BMI (median (IQR) 26.7kg/m^2 (24.2 to 30.1) vs 26.9kg/m^2 (24.1 to 30.5)) and waist circumference (median (IQR) 81 cm (74 to 90) in both groups). Also, no notable differences were observed for the distribution of other variables between the two groups (Supplementary Table B-3).

Results from complete case analysis are provided for all mediation analyses (Table 6-5 and Supplementary Table B-2). Compared with multiple imputation analyses, for both BMI and waist circumference, the TEs estimated from complete case analyses were weaker, while the IIEs through calculated free estradiol were stronger. There was no evidence for mediating effect through fasting insulin using the complete case analysis.

Discussion

We estimated that increased calculated free estradiol concentrations explain approximately half of the association between measures of adiposity (BMI or waist circumference) and ER-positive breast cancer in postmenopausal women. There was no evidence of a mediating effect through insulin.

To our knowledge, of the existing studies that have quantified the mediating effect of estradiol on adiposity-postmenopausal breast cancer association (120, 135), our study is the only study with a measure for free estradiol. A strength of mediation analysis based on interventional effects is that it allowed us to quantify the mediating effects of fasting insulin and calculated free estradiol without assuming that the two mediators do not influence each other directly (130). The method also allowed for inclusion of possible exposure-mediator and mediator-mediator interactions in the models (130, 131). These

are important advantages over the traditional approaches used to assess mediation, which are invalid in the presence of exposure-mediator interactions or multiple interrelated mediators (155). While we controlled for the known measured confounders of exposure-outcome, mediator-outcome, and exposure-mediator relationships, we did not measure biomarkers in other biological mechanisms that could also serve as mediators of the obesity-breast cancer relationship such as the inflammatory pathway. Inflammatory factors might influence fasting insulin and free estradiol levels, as well as postmenopausal breast cancer risk (163). Ignoring these potential mediator-outcome confounders might have led to overestimation of the indirect effect through calculated free estradiol and fasting insulin. Our results might have also been biased due to residual confounding from unmeasured (e.g. family history of breast cancer) or imperfectly measured confounders. Another limitation of our study was the lack of longitudinal measurements for adiposity and biomarkers (we used measures collected at baseline), which precluded us from ruling out reverse causation with respect to the exposure and mediator. Preclinical cancer might have influenced adiposity or levels of the biomarkers; to address this, we excluded all cases diagnosed within the first year after recruitment.

BMI and waist circumference were measured following standard protocols, and breast cancer cases were ascertained through linkage to a population-based cancer registry, minimising measurement error. The limitations of using BMI to define adiposity are well known (13). Using the percentage of body fat measured by dual-energy x-ray absorptiometry scans for validation, it has been observed that, depending on age and race, BMI has a sensitivity ranging from 42.1% (95% CI 36.2% to 48.1%) to 73.7% (95% CI 65.5% to 82.8%) to identify postmenopausal obese women (177). The non-differential misclassification of obesity based on BMI (i.e. misclassification of exposure independent of outcome) might have led us to underestimate the BMI-breast cancer association in our

analyses. We also report results for waist circumference as another proxy for adiposity, which is considered a better indicator of active visceral adipose tissue and the preferred measure of adiposity (13). Circulating total estradiol and SHBG, which were used in calculation of free estradiol, were well-measured and stable over time, as demonstrated by high ICCs in the reliability study and/or low coefficient of variation for the assays (138). Sex-steroid hormones levels in serum and breast adipose tissue are highly correlated (166, 178) and it is likely that circulating hormones levels are reflective of hormone concentration in breast tissue (166). We did not have measures of reliability for the assay used to measure fasting insulin. We had measures for insulin-like growth factor-1 and binding protein-3 in our cohort but did not include these as potential mediators. Although there is strong evidence supporting a positive association between these biomarkers and postmenopausal breast cancer risk(179), their relation with obesity is uncertain (173). Our study excluded women who, at baseline, had a history of diabetes, because insulin production might decrease in response to long-term hyperglycaemia (180). This exclusion might have led to an underestimation of the indirect effect through insulin. Perhaps the most important limitation of our study was that biomarker measurements were not available for 37% of the women because they were either not fasting at the time of blood collection (28%) or had insufficient or contaminated plasma. This missingness was handled using multiple imputation and assuming that the data were missing dependent only on measured variables.

Results of a causal mediation analysis from the WHI-OS suggested that of the overall BMI- postmenopausal breast cancer association (additional cases per 100,000 women for a 5 kg/m^2 increment in BMI 50.0 (95% CI 23.2 to 76.6)), 24% could be attributed to total estradiol (indirect additional cases per 100,000 11.9 (95% CI 1.7 to 22.6)) (120). The proportion mediated through total estradiol was higher for ER-positive cancers (49%). A

sub-study from the PLCO, reported the estimated total effect for a 5 kg/m^2 increment in BMI on risk of ER-positive postmenopausal breast cancer as 7.0 (95% CI 2.1 to 11.9) additional cases per 100,000, of which 20% was through total estradiol (indirect additional cases per 100,000 1.4 (95% CI -1.0,3.9)) (135). In our study, 51% of the TE of BMI per 5 kg/m^2 and 30% of the TE of waist circumference per 10 cm on ER positive breast cancer could be attributed to calculated free estradiol. The stronger indirect effect through free estradiol estimated in our study compared with the reported indirect effects through total estradiol in the other two studies is in line with the evidence suggesting that bioavailable estradiol is a more important player in explaining the adiposity-breast cancer association than total estradiol (121, 167). Together, these observations imply that estradiol, particularly free estradiol, plays a substantial role in mediating the effects of adiposity on breast cancer risk. But despite its importance, free estradiol does not explain all the association between adiposity and breast cancer.

In the WHI-OS, about 66% of the BMI-postmenopausal breast cancer association in non-diabetic women was mediated through insulin (additional cases per 100,000 for the overall effect for 5 kg/m^2 increment in BMI 52.0 (95% CI 12.1 – 91.3) and indirect effect through insulin 34.2 (95% CI 9.4 – 59.0)) (120). In our study, there was no evidence for an indirect effect through insulin. In line with this difference, the observed association between fasting insulin and postmenopausal breast cancer was weaker in our study compared with that reported for the WHI-OS (HR for postmenopausal breast cancer in non-hormone therapy users for highest vs. lowest quartiles of fasting insulin 2.40 (95% CI 1.30 to 4.41)) (119). Different assays used to measure insulin might explain some of the difference in the insulin-postmenopausal breast cancer association observed in our study and WHI-OS (91). In a meta-analysis investigating the association between serum insulin and breast cancer, the summary relative risk for the association between insulin

and breast cancer was 1.02 (95% CI 0.53 to 1.97) for six prospective studies. The association did not change when analysis was limited to studies with postmenopausal breast cancer as the outcome (91). A high heterogeneity was reported across studies (I^2 73%), which was suggested to be in part explained by the variety of assays used to measure insulin levels (91). Further research is required to examine the presence and magnitude of a mediating role of the insulin signalling pathway in the adiposity-breast cancer association.

In summary, we applied novel mediation methods to quantify the mediating effect of two potentially interrelated pathways on the association between adiposity and ER-positive postmenopausal breast cancer. Our results provide additional evidence that free estradiol plays a key role in mediating the adiposity-postmenopausal breast cancer relationship. However other pathways must also be important for explaining this association; identifying other contributing pathways and quantifying their contribution to breast cancer development should be a priority.

Acknowledgements: The MCCS was made possible by the contribution of many people, including the original investigators, the teams that recruited the participants and continue working on follow-up, and the many thousands of Melbourne residents who continue to participate in the study. Where authors are identified as personnel of the International Agency for Research on Cancer / World Health Organization, the authors alone are responsible for the views expressed in this article and they do not necessarily represent the decisions, policy or views of the International Agency for Research on Cancer / World Health Organization.

Financial support: This work was supported by Australian National Health and Medical Research Council grants (209057, 396414, 504711, and 1150591 to SGD and 1104975 to JAS); National Breast Cancer Foundation Fellowship (ECF-15-012 to BML); Melbourne Research Scholarship (to SGD); and by infrastructure provided by Cancer Council Victoria. The Melbourne Collaborative Cohort Study (MCCS) cohort recruitment was funded by VicHealth and Cancer Council Victoria. Cases and their vital status were ascertained through the Victorian Cancer Registry and the Australian Institute of Health and Welfare, including the National Death Index and the Australian Cancer Database. The work reported in this paper was undertaken by SGD while hosted by the International Agency for Research on Cancer (IARC) and funded partially by a PhD Scholarship by IARC.

Data availability: The data that support the findings of this study are held by the Cancer Epidemiology Division, Cancer Council Victoria. Restrictions apply to the availability of these data, which were used under license for this study. Information on how data from the Melbourne Collaborative Cohort Study (Health 2020) data can be accessed is available at https://www.pedigree.org.au/.

Conflict of interest: The authors declare no potential conflicts of interest.

Chapter 7: Adiposity and endometrial cancer risk in postmenopausal women in the EPIC

In this chapter, an analysis of a case-control study nested within the European Prospective Investigation into Cancer and Nutrition (EPIC) has been carried out to investigate the mediating roles of adiponectin, interleukin-6 (IL-6), interleukin-1-receptor antagonist (IL-1Ra), TNF-receptor-1 (TNF-R1), TNF-R2, and CRP (to reflect the inflammatory status), C-peptide (to reflect hyperinsulinemia), and calculated free estradiol and estrone (to reflect alterations in sex-steroid hormone levels) in the effect of adiposity on endometrial cancer risk in postmenopausal women. This chapter includes the manuscript submitted to Cancer Epidemiology, Biomarkers & Prevention. Maintaining the confidentiality of the findings in this paper, required by the Journal, is, therefore, kindly requested.

For consistency, spelling, referencing, tables, and figure numbers have been modified. Supplementary material related to this publication are presented in Appendix C.

Adiposity and endometrial cancer risk in postmenopausal women: a sequential causal mediation analysis

S Ghazaleh Dashti,[1,2] Dallas R English,[1,3] Julie A Simpson,[1] Amalia Karahalios,[1,4] Margarita Moreno-Betancur, [5,6] Carine Biessy,[2] Sabina Rinaldi,[2] Pietro Ferrari,[2] Anne Tjønneland,[7] Jytte Halkjær,[7] Christina C Dahm,[8] Helene Tilma Vistisen,[8] Florence Menegaux,[9] Vittorio Perduca,[10] Gianluca Severi,[9,11,12] Krasimira Aleksandrova,[13,14] Matthias B. Schulze,[14,15] Giovanna Masala,[16] Sabina Sieri,[17] Rosario Tumino,[18] Alessandra Macciotta,[19] Salvatore Panico,[20] Anouk E. Hiensch,[21] Anne M. Ma,[21] J. Ramón Quirós,[22] Antonio Agudo,[23] Maria-Jose Sánchez,[24,25,26,27] Pilar Amiano,[28,29] Sandra Colorado-Yohar,[30,31,32] Eva Ardanaz,[33,34,35] Naomi E Allen,[36] Elisabete Weiderpass,[37] Renée Turzanski Fortner,[38] Sofia Christakoudi,[39,40], Konstantinos K. Tsilidis,[39,41] Elio Riboli,[39] Rudolf Kaaks,[38] Marc J Gunter,[2] Vivian Viallon,*[2] Laure Dossus,*[2]

*Joint senior authors

1 Centre for Epidemiology and Biostatistics, Melbourne School of Population and Global Health, The University of Melbourne, Melbourne, Victoria, Australia
2 Section of Nutrition and Metabolism, International Agency for Research on Cancer (IARC), Lyon, France
3 Cancer Epidemiology and Intelligence Division, Cancer Council Victoria, Melbourne, Victoria, Australia
4 School of Public Health and Preventive Medicine, Monash University, Melbourne, Victoria, Australia
5 Department of Paediatrics, The University of Melbourne, Melbourne, Victoria, Australia
6 Clinical Epidemiology and Biostatistics Unit, Murdoch Children's Research Institute, Melbourne, Australia
7 Danish Cancer Society Research Center, Copenhagen Denmark
8 Department of Public Health, Aarhus University, Aarhus, Denmark
9 Université Paris-Saclay, UVSQ, CESP U1018 INSERM, Villejuif, France
10 Laboratoire de Mathématiques Appliquées à Paris 5—MAP5 (UMR CNRS 8145), Université Paris Descartes, Université de Paris, Paris, France
11 Gustave Roussy, Villejuif, France
12 Department of Statistics, Computer Science, Applications “G. Parenti”, University of Florence, Italy
13 Nutrition, Immunity and Metabolism Senior Scientist Group, Department of Nutrition and Gerontology, German Institute of Human Nutrition, Potsdam-Rehbruecke (DIfE), Nuthetal, Germany
14 University of Potsdam, Institute of Nutritional Science, Potsdam, German
15 Department of Moleculr Epidemiology, German Institute of Human Nutrition Potsdam-Rehbruecke, Nuthetal, German
16 Cancer Risk Factors and Life-Style Epidemiology Unit, Institute for Cancer Research, Prevention and Clinical Network - ISPRO, Florence, Italy
17 Epidemiology and Prevention Unit, Fondazione IRCCS Istituto Nazionale dei Tumori di Milano
18 Cancer Registry and Histopathology Department, Provincial Health Authority (ASP) Ragusa, Italy
19 Department of Clinical and Biological Sciences, University of Turin, Italy
20 Dipoartimento Di Medicina Clinica E Chirurgia Federico II University, Naples, Italy
21 Julius Center for Health Sciences and Primary Care, University Medical Center Utrecht, Utrecht University, Utrecht, The Netherlands
22 Public Health Directorate, Asturias, Spain
23 Unit of Nutrition and Cancer, Catalan Institute of Oncology - ICO, Nutrition and Cancer Group, Bellvitge Biomedical Research Institute - IDIBELL, L'Hospitalet de Llobregat, Barcelona, Spain
24 Escuela Andaluza de Salud Pública (EASP), Granada, Spain
25 Instituto de Investigación Biosanitaria ibs.GRANADA, Granada, Spain
26 Centro de Investigación Biomédica en Red de Epidemiología y Salud Pública (CIBERESP), Madrid, Spain
27 Universidad de Granada, Granada, Spain

28 Ministry of Health of the Basque Government, Public Health Division of Gipuzkoa, Biodonostia Health Research Institute, Donostia-San Sebastian, Spain
29 CIBER Epidemiología y Salud Pública, Madrid, Spain
30 Department of Epidemiology, Murcia Regional Health Council, IMIB-Arrixaca, Murcia, Spain
31 CIBER Epidemiología y Salud Pública (CIBERESP), Spain.
32 Research Group on Demography and Health, National Faculty of Public Health, University of Antioquia, Medellín, Colombia
33 Navarra Public Health Institute, Pamplona, Spain
34 IdiSNA, Navarra Institute for Health Research, Pamplona, Spain
35 CIBER Epidemiology and Public Health CIBERESP, Madrid, Spain
36 Nuffield Department of Population Health, University of Oxford, Oxford, United Kingdom
37 International Agency for Research on Cancer (IARC), Lyon, France
38 Division of Cancer Epidemiology, German Cancer Research Center (DKFZ), Heidelberg, Germany
39 Department of Epidemiology and Biostatistics, School of Public Health, Imperial College London, Norfolk Place, St Mary's Campus, London W2 1PG, United Kingdom
40 MRC Centre for Transplantation, King's College London, Great Maze Pond, London SE1 9RT, United Kingdom
41 Department of Hygiene and Epidemiology, University of Ioannina School of Medicine, Ioannina, Greece

Corresponding author:
S Ghazaleh Dashti
Centre for Epidemiology and Biostatistics
Melbourne School of Population and Global Health
Level 3, 207 Bouverie Street
The University of Melbourne VIC 3010 Australia
Email: seyedeh.dashti@unimelb.edu.au

Conflict of interest: The authors declare no potential conflicts of interest.

Abbreviations: EC endometrial cancer; IGF-1 insulin-like growth factor-1; SHBG sex hormone binding globulin; EPIC European Prospective Investigation into Cancer and Nutrition; IL-6 interleukin-6; IL-1Ra interleukin-1 receptor antagonist, TNF-α tumour necrosis factor-a; TNFR-1 TNF receptor 1, TNFR-2 TNF receptor 2, CRP C-reactive protein; BMI body mass index; GMR Geometric mean ratios; CI confidence interval; OR odds ratio; TE total effect; NIE natural indirect effect; NDE nature direct effect; PM proportion mediated; CV coefficient of variation; IGFBP insulin-like growth factor binding protein; TyG triglyceride glucose product; HR hazard ratio

Abstract

Background

Adiposity increases endometrial cancer (EC) risk, possibly through inducing inflammation, hyperinsulinemia, and increasing estrogens. We aimed to quantify how much of the adiposity-EC link in postmenopausal women was mediated by adiponectin (anti-inflammatory adipocytokine); interleukin-6, interleukin-1-receptor antagonist, TNF-receptor-1 and -2, and C-reactive protein (biomarkers of inflammatory status); C-peptide (hyperinsulinemia biomarker); free estradiol and estrone (estrogen biomarkers).

Methods

We used data from a case-control study nested within the European Prospective Investigation into Cancer and Nutrition. Eligible women were those who at recruitment had not had cancer, hysterectomy, diabetes, did not use oral contraceptives or hormone therapy, were postmenopausal, and had no missing confounder data. Multiple mediating pathways from adiposity measures to EC were investigated by estimating natural indirect (NIE) and direct (NDE) effects using sequential mediation analysis.

Results

The study included 163 cases and 306 controls. The adjusted odds ratio (OR) for EC for BMI≥30 *vs.* 18.5≤BMI<25kg/m^2 was 2.51 (95%CI 1.26–5.02). The ORs^{NIE} were 1.95 (1.01–3.74) through all biomarkers (72% proportion mediated (PM) on log(OR) scale) decomposed as: 1.35 (1.06–1.73) through pathways originating with adiponectin (33%PM); 1.13 (0.71–1.80) through inflammation beyond [the potential influence of] adiponectin (13%PM); 1.05 (0.88–1.24) through C-peptide beyond adiponectin and inflammation (5%PM); and 1.22 (0.89–1.67) through estrogens beyond previous

biomarkers (21%PM). The OR^{NDE} not through biomarkers was 1.29 (0.54–3.09). Waist circumference gave similar results.

Conclusion

Reduced adiponectin and increased inflammatory biomarkers, C-peptide, and estrogens mediated about 70% of increased odds of EC in women with obesity vs normal weight.

Impact

If replicated, these results could have implication for identifying targets for intervention to reduce EC risk in women with obesity.

Introduction

Excess adiposity is an important risk factor for endometrial cancer (EC) (181). In 2012, 34% (90% confidence interval (CI) 32%–36%) of diagnosed EC cases were attributable to overweight and obesity (3). Disturbed adipocytokine production and inflammation, hyperinsulinemia, and sex-steroid hormones – are hypothesized to underly the adiposity-EC link (182). In women with obesity, adipose tissue secretes less adiponectin (an anti-inflammatory adipocytokine) and more inflammatory adipocytokines (183). This inflammatory status may have mitogenic, anti-apoptotic, and angiogenic effects (5, 6); reduce insulin sensitivity (8); or dysregulate aromatase expression and increase estrogen levels (54). Insulin sensitivity may also be reduced through increased hepatic glucose production in response to excess free fatty acids (8). The resulting hyperinsulinemia and higher circulating insulin levels is causally linked to EC (49). It may have mitogenic effects; increase free insulin-like growth factor-1 (IGF-1) levels (14, 15); increase aromatase activity via IGF-1, thus increase estrogen levels (54); or down-regulate sex-hormone-binding globulin (SHBG) production and increase bioavailable estrogens (184). Bioavailable estrogens, especially when unopposed by progesterone, (182) may increase EC risk through mitogenic effects in endometrial tissue (173).

Causal mediation analysis (155) can quantify the role of biological pathways involved in the effect of adiposity on EC risk. As highlighted, it is improbable that the influences of the involved biomarkers are siloed. When measures for multiple biomarkers are available, mediation analysis approaches should account for correlations between biomarkers (155). Failing to do so and assessing the mediating roles of biomarkers

individually may result in biased estimation of the effect explained by the biomarkers (155).

We used data from a case-control study nested within the European Prospective Investigation into Cancer and Nutrition (EPIC) (118, 139) to quantify the mediating roles of biomarkers representing inflammatory status, hyperinsulinemia, and estrogens in explaining the effect of adiposity on EC risk in postmenopausal women. Although adiposity increases pre- and postmenopausal EC risk (181), we excluded premenopausal women because the mechanisms may be different before and after menopause (185).

Methods

Established in 1992, EPIC is a cohort study including ~370,000 women recruited from ten European countries. Details of baseline data collection, blood sample collection and storage, ascertainment of cancers and follow-up are published (139). The present study made use of data from a nested case-control study that investigated associations between EC risk and sex-steroid hormones, metabolic factors, and inflammatory biomarkers measured in baseline blood samples, for which follow-up ended between December 1999 and November 2003 (118).

Selection of cases and controls

The eligibility criteria for the original case-control study included no history of hysterectomy or diagnosis of cancer (except keratinocyte skin cancer) and no use of oral contraceptives or hormone therapy at blood collection. It included 233 incident primary EC cases and 446 controls (incidence density sampling), matched on recruitment centre, age and fasting status at blood draw, time of blood draw, menopausal status (118). One case was later found to have prevalent cancer at recruitment, and seven were no longer

classified as EC cases. We excluded these women and their matched controls. Additionally, women who at recruitment were not postmenopausal (46 cases, 86 controls), had a history of diabetes (9 cases, 16 controls), had BMI<18.5kg/m^2 (1 case, 3 controls), or had missing values for the confounders (6 cases, 17 controls) were excluded from the analysis.

Biomarkers

Measured biomarkers were adiponectin; interleukin-6 (IL-6), interleukin-1 receptor antagonist (IL-1Ra), tumor necrosis factor-α (TNF-α), TNF receptor 1 (TNFR-1), TNFR-2, and C-reactive protein (CRP) (biomarkers of inflammatory status); C-peptide (hyperinsulinemia biomarker); calculated free estradiol (proxy for bioavailable estradiol) and estrone (estrogen pathway biomarkers) (77, 78, 94, 113, 140). Free estradiol was calculated from SHBG and total estradiol (186).

Statistical Analyses

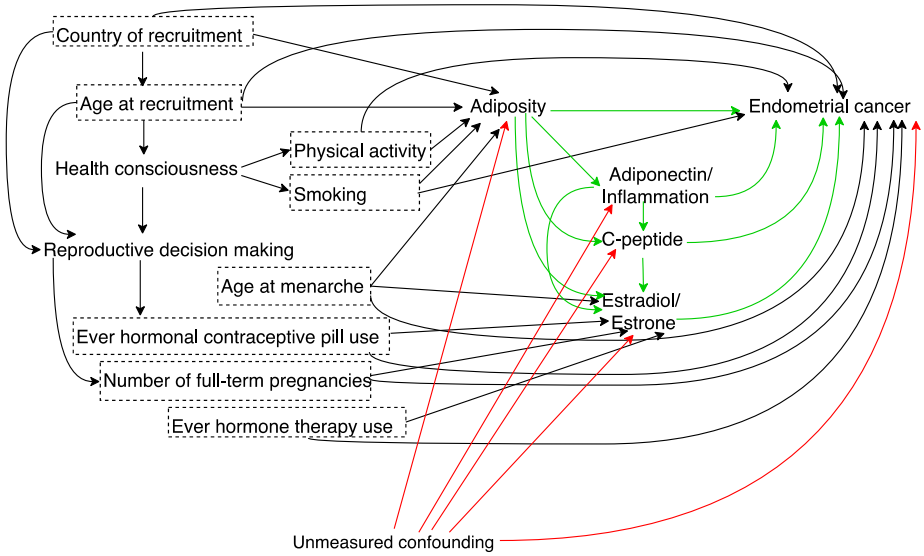
We had three measures of adiposity: BMI, waist circumference, and waist-hip ratio. To assist with the interpretations of the mediated effect and translation to policy and to provide the greatest contrasts, in primary analyses these were included as binary variables comparing women with obesity *vs.* normal. The cut-off values to categorize women were based on guidelines (BMI ≥30*vs.* ≥18.5-<25 kg/m^2; waist circumference >88*vs.*≤80 cm) or tertile cut-offs (waist-hip ratio >0.84*vs.*≤0.78) (25).

To remove the effects of batch, fasting status and time of blood draw, for each biomarker we (i) fitted a linear mixed-effects model to the log-transformed biomarker value for the controls with batch as a random effect, fasting status at blood draw and time of blood draw as fixed effects; then (ii) for all women, derived a normalized value

by subtracting the difference between the predicted mean batch-specific values and the overall mean from the observed values.

***Confounder selection**:* The sequential mediation analysis used (see *Mediation analysis*) relied on no unmeasured exposure-outcome, mediator-outcome, and exposure-mediator confounding. A causal diagram (Figure 7-1) was developed with reference to existing evidence (45, 173, 187, 188) to identify these confounders (155). These potential confounders included age at recruitment, number of full-term pregnancies, age at menarche and menopause, history of oral contraceptive use and hormone therapy, smoking status, and physical activity. All variables had <3% missing values, except age at menopause (7% missing), which was not included because it was weakly correlated with the exposure and biomarkers ($r<|0.17|$).

Figure 7-1 Assumed causal structure underlying the effect of adiposity on endometrial cancer



To avoid overloading the diagram, we did not include all the possible arrows between all the variables. We only included the arrows that were sufficient to flag a variable as a common cause (confounder) of either exposure-outcome, mediator-outcome, or exposure-mediator associations. Therefore, the diagram is not strictly a "causal diagram". The green arrows represent pathways (indirect and direct) that we were interested in. The red arrows represent the potentially biasing paths due to unmeasured confounding. Variables in the dashed boxed were included (conditioned on) in all multivariable analyses. Based on the assumed causal structure represented in this diagram, conditioning on none of these variables would have introduced collider bias.

Exposure-mediator associations: Geometric mean ratios (GMR) of biomarkers (and 95% CIs) in relation to adiposity measures were estimated using linear regression models applied to log-transformed data for the controls, with adjustment for potential confounders.

Exposure-outcome and mediator-outcome associations*:* In models that included the outcome, we broke the matching to avoid losing participants and included the matching variables as covariates. We grouped the 21 recruitment cente into regions to avoid creating combinations with sparse data. Age at recruitment was modelled as restricted cubic splines. The ORs and 95% CIs for the association between EC and adiposity measures and biomarkers were estimated from models that included confounders. Models for biomarkers included BMI and when applicable, other biomarkers that might have confounded the biomarker-outcome association. The linearity of mediator-outcome

associations was checked using models with restricted cubic splines (2 degrees of freedom).

Mediation analysis with multiple mediators: Figure 7-1 guided the analysis. The association between each adiposity measure and EC (total effect (TE) (155)) was decomposed into a natural indirect effect (NIE) through all biomarkers and a natural direct effect (NDE) that captured the remaining association not through any of the biomarkers (see Figure 7-2 and reference (155) for definitions of the NIE and NDE). Assuming that reduced adiponectin and increased inflammation biomarkers levels preceded and potentially, but not necessarily, influenced C-peptide and estrogen levels, and C-peptide preceded and potentially influenced estrogens (8, 54, 149) we sequentially (129) decomposed the estimated NIE into NIEs through 1) pathways originating with reduced adiponectin and increased inflammation biomarkers; 2) C-peptide beyond the influences of adiponectin and inflammation biomarkers and; 3) and estrogens (free estradiol and estrone), beyond the influences of adiponectin, inflammation biomarkers and C-peptide. Assuming that adiponectin preceded and influenced inflammation biomarkers (149), we additionally decomposed the first NIE into NIEs through 1) pathways originating with reduced adiponectin and 2) inflammation biomarkers beyond the potential influence of adiponectin. Figure 7-2 shows interpretations of these effects.

The NIEs and NDE were estimated on the log(OR) scale using a regression-standardisation approach, with linear regression models for the mediators and logistic regression models for the outcome (see Supplementary Table C-1 for details) (129). To take the matched design into account, mediator models were limited to controls and the TE, NDE, and NIEs were estimated with the effects standardized to the confounder

distribution for all eligible women in EPIC (128). The proportion mediated (PM) was calculated on the log(OR) scale as $\log(OR)^{NIE}/(\log(OR)^{NDE} + \log(OR)^{NIE})$) (155).

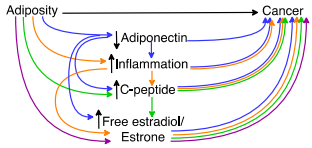
Missing data: Women with missing data for confounders were excluded because they comprised a small proportion (5%) of the participants. Seventeen percent of women had missing biomarker data, which was multiply imputed based on chained equations with 20 iterations (175). The imputation models included all variables in the mediation analyses, and recruitment centre, height, weight, and hip circumference as auxiliary variables. Within each imputed dataset, 1,000 bootstrap samples were used to estimate standard errors for the TE, NDE, and NIEs. These estimates were pooled using Rubin's rules to calculate the final estimates and 95% CI (146).

Sensitivity Analysis: We also repeated the mediation analyses with continuous adiposity measures, and after excluding cases diagnosed within two years post-recruitment.

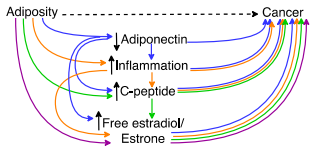
All analyses were performed in Stata version 15.1 (189).

Figure 7-2 Interpretation of natural indirect and direct effects; the solid arrows represent pathways estimated under each effect.*

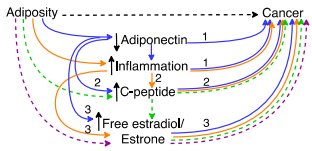
1) The total effect
The effect of adiposity on endometrial cancer risk exerted through adiposity-induced changes in levels of all biomarkers of interest (natural indirect effect) and other possible pathways not captured by the included biomarkers (natural direct effect).



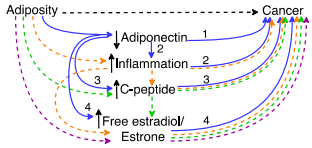
2) The indirect effect through all the biomarkers
The effect of adiposity on endometrial cancer risk exerted through adiposity-induced changes in levels of all biomarkers of interest. This effect does not rely on any assumptions about the causal ordering of the biomarkers.



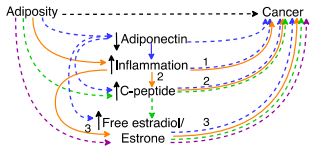
3) The indirect effect through reduced adiponectin and increased inflammation**
The effect the adiposity-induced reduction in adiponectin levels and increased inflammation biomarkers levels might have on endometrial cancer risk either directly (1), or by further influencing the levels of C-peptide (2), or estrogens (3). This effect does not rely on any assumptions about the causal ordering between reduced adiponectin and increased inflammation biomarkers levels.



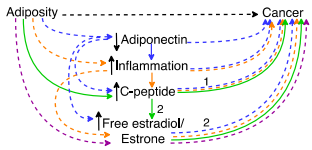
4) The indirect effect through reduced adiponectin levels
The effect adiposity-induced reduction in adiponectin levels might have on endometrial cancer risk either directly (1), or by further influencing the levels of inflammatory biomarkers (2), C-peptide (3), or estrogens (4). This effect assumes that adiposity-induced reduction in adiponectin levels precedes and potentially influences inflammation biomarker levels and not vice versa.



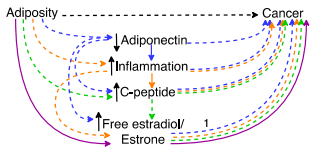
5) The indirect effect through increased inflammation beyond the potential influence of adiponectin
The effect adiposity-induced increased inflammation biomarkers levels might have on endometrial cancer risk either directly (1), or by further increasing C-peptide (2), or estrogen levels (3). This effect excludes the potential influence of adiponectin on inflammation biomarkers. Thus, similar to the effect #4, it assumes that adiposity-induced reduction in adiponectin levels precedes and potentially influences inflammation biomarker levels and not vice versa.



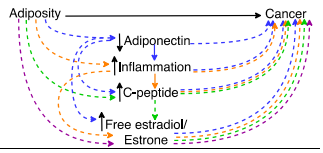
6) The indirect effect through increased C-peptide beyond the potential influences of adiponectin and inflammation
The effect adiposity-induced increased C-peptide levels might have on endometrial cancer risk either directly (1), or by further increasing estrogens levels (2). This effect excludes the potential influences of adiponectin and inflammation biomarkers on C-peptide levels.



7) The indirect effect through estrogens beyond the potential influences of adiponectin, inflammation, and c-peptide
The effect adiposity-induced increased estrogens levels might have on endometrial cancer risk (1). This effect excludes the potential influences of adiponectin, inflammation biomarkers, and C-peptide on estrogen levels.



8) The direct effect not through any of the biomarkers
The effect of adiposity on endometrial cancer risk through pathways not captured by the included biomarkers



*In the terminology of causal mediation, the natural indirect effect is defined as the average change in the outcome, if the exposure status (e.g. adiposity) was set to what it would be if everybody was exposed (having obesity), and the mediator was allowed to change from what it would naturally be if exposed (with obesity) to what it would naturally be if unexposed (with normal weight). The natural direct effect is the average change in the outcome, if the exposure status (e.g. adiposity) was allowed to change from exposed (with obesity) to unexposed (with normal weight) and the mediator level was set to what it would naturally be if unexposed (with normal weight).

Results

The analytic dataset included 163 EC cases and 306 controls. Median age at EC diagnosis was 63 years (interquartile range 60–68) (Table 7-1). At baseline, compared with controls, a smaller proportion of cases had used oral contraceptives, were current smokers, had moderate/high physical activity; a larger proportion had used hormone therapy, had BMI≥30 kg/m^2 or waist circumference>88 cm. Compared with women with no missing data, a higher proportion of women with missing biomarker values were cases, fasting >6 hours at blood draw, from Northern Europe, current smokers, or with BMI≥30 kg/m^2 (Supplementary Table C-2).

Table 7-1 Baseline characteristics of endometrial cancer cases and controls

	Controls N=306	Cases N=163
Age at blood collection, *years*; median (interquartile range)	60.0 (56.6-63.0)	60.4 (56.7-63.2)
Fasting status at blood draw, *hours*; n(%)		
<3 hours	154 (50)	85 (52)
3-6	56 (18)	27 (17)
>6	96 (31)	51 (31)
Region of recruitment*; n(%)		
Western Europe	77 (25)	43 (26)
Northern Europe	119 (39)	66 (40)
Southern Europe	110 (36)	54 (33)
Full-term pregnancies, *number*; median (interquartile range)	2.0 (1.0-3.0)	2.0 (1.0-3.0)
Age at menarche, *years;* median (interquartile range)	13.0 (12.0-14.0)	13.0 (12.0-14.0)
Had history of oral contraceptive use; n(%)	109 (36)	47 (29)
Had history of hormone therapy; n(%)	47 (15)	36 (22)
Smoking status, n(%)		
never	196 (64)	112 (69)
former	54 (18)	33 (20)
current	56 (18)	18 (11)
Physical activity; n(%)		
Inactive to moderately inactive	106 (35)	64 (39)
moderately active to active	200 (65)	99 (61)
Body Mass Index, kg/m^2 ; n(%)		
<25	129 (42)	47 (29)
≥25 & <30	121 (40)	61 (37)
≥30	56 (18)	55 (34)
median (IQR)	25.9 (23.5-28.5)	27.5 (24.0-32.4)
Waist Circumference, *cm*; n(%)		
≤80	122 (40)	54 (33)
>80 & ≤ 88	94 (31)	31 (19)
>88	90 (29)	78 (48)
median (interquartile range)	83.0 (76.0-91.0)	88.0 (78.5-95.5)
Waist-hip ratio; n(%)		
≤0.78	106 (35)	47 (29)
>0.78 & ≤0.84	110 (36)	56 (34)
>0.84	90 (29)	60 (37)
median (interquartile range)	0.8 (0.8-0.9)	0.8 (0.8-0.9)
Biomarkers; median (interquartile range)		
Adiponectin, *μg/mL*	10.87 (7.84-14.21)	8.94 (5.97-12.20)
Adiponectin missing value	1 (0)	0 (0)
Interleukin-6, *pg/mL*	1.2 (0.9-2.0)	1.4 (1.0-2.3)
Interleukin-6 missing value	11 (4)	9 (6)
Interleukin-1 receptor antagonist, *pg/mL*	22.4 (18.2-64.4)	35.6 (18.4-141.3)
Interleukin-1 receptor antagonist missing value	5 (2)	2 (1)
Tumour necrosis factor-*α*, *pg/mL*	1.0 (0.6-1.4)	1.0 (0.7-1.6)
Tumour necrosis factor-*α* missing value	3 (1)	1 (1)
Tumour necrosis factor-receptor 1, *pg/mL*	998.0 (912.1-1151.8)	1075.6 (920.0-1233.8)
Tumour necrosis factor-receptor 1 missing value	2 (1)	0 (0)
Tumour necrosis factor-receptor 2, *pg/mL*	1909.5 (1676.7-2187.8)	1988.4 (1722.1-2388.6)
Tumour necrosis factor-receptor 2 missing value	2 (1)	0 (0)
C-reactive protein, *ng/mL*	1345.6 (691.9-2640.7)	1745.1 (989.0-3223.7)
C-reactive protein missing value	7 (2)	8 (5)
C-peptide, *ng/mL*	3.1 (2.3-4.1)	3.4 (2.5-4.9)
C-peptide missing value	0 (0)	0 (0)
Calculated free estradiol, *pg/mL*	2.0 (1.6-2.5)	2.3 (1.8-3.2)
Free estradiol missing value	6 (2)	5 (3)
Estrone, *pg/mL*	32.7 (25.9-39.1)	35.7 (29.6-46.6)
Estrone missing value	23 (8)	15 (9)
Missing value for any of the biomarkers	47 (15)	32 (20)

* Centres from France, Germany, and the Netherlands were categorized as Western Europe; centres from Denmark, and the United Kingdom as Northern Europe; and centres from Italy, Spain, and Greece as Southern Europe.

Exposure-mediator associations: For BMI (≥30 *vs*. ≥18.5-<25 kg/m^2) positive associations were observed with IL-6, IL-1Ra, TNF-R1, TNF-R2, CRP, C-peptide, free estradiol and estrone. Of these, CRP demonstrated the strongest association (GMR 2.80; 95%CI 2.06–3.79). An inverse association was observed for adiponectin (GMR 0.77; 95%CI 0.67–0.87). Similar patterns were seen for waist circumference and waist-hip ratio (Table 7-2; Supplementary Table C-3 complete-case analysis results).

Table 7-2 Exposure-mediator association; estimated ratios in geometric mean of biomarkers associated with body mass index and waist circumference; multiple imputation analyses limited to controls and excluding women categorized as overweight

	Ratio of geometric means (95% confidence interval)								
	Body Mass Index ≥30 vs. <25 kg/m2			Waist circumference >88 vs. ≤80 cm			Waist-hip ratio >0.84 vs. ≤0.78 cm		
Biomarker	n=185			n=212			n=196		
Adiponectin	0.77	(0.67 to	0.87)	0.69	(0.61 to	0.78)	0.66	(0.59 to	0.75)
Interleukin-6	1.87	(1.52 to	2.30)	1.67	(1.41 to	1.98)	1.63	(1.34 to	1.98)
Interleukin-1 receptor antagonist	1.55	(1.09 to	2.20)	1.50	(1.14 to	1.99)	1.51	(1.10 to	2.06)
Tumour necrosis factor-α	1.06	(0.85 to	1.32)	1.06	(0.88 to	1.27)	1.14	(0.95 to	1.38)
Tumour necrosis factor-receptor 1	1.21	(1.14 to	1.30)	1.17	(1.10 to	1.24)	1.10	(1.03 to	1.17)
Tumour necrosis factor-receptor 2	1.19	(1.11 to	1.28)	1.12	(1.05 to	1.20)	1.06	(0.99 to	1.14)
C-reactive protein	2.80	(2.06 to	3.79)	2.52	(1.96 to	3.26)	2.38	(1.79 to	3.15)
C-peptide	1.48	(1.26 to	1.74)	1.57	(1.38 to	1.79)	1.52	(1.32 to	1.74)
Free estradiol	1.28	(1.13 to	1.45)	1.35	(1.21 to	1.50)	1.30	(1.16 to	1.46)
Estrone	1.16	(0.99 to	1.35)	1.14	(1.00 to	1.29)	1.09	(0.97 to	1.23)

Models were adjusted for age at blood collection, time at blood collection, fasting status at blood collection, recruitment centre, age at menarche, number of full-term pregnancies, history of oral contraceptive use, history of hormone therapy, physical activity, and smoking status.

Exposure-outcome and mediator-outcome associations: An increased OR for EC was observed for BMI ≥30 *vs*. ≥18.5-<25 kg/m^2 (OR 3.00; 95%CI 1.70–5.28) and waist circumference >88 *vs*. ≤80 cm (OR 2.08; 95%CI 1.29–3.35). The evidence for an association between waist-hip ratio and EC risk was weak (OR >0.84 *vs*. ≤0.78 1.57; 95% CI, 0.93-2.94) (Table 7-3). Therefore, we did not attempt to decompose this effect in our mediation analysis.

An inverse association was observed between adiponectin and EC, (OR per doubling concentration 0.65; 95%CI 0.47–0.90), and a positive association for IL-1Ra (OR 1.14;

95%CI 1.00 to 1.29), and estrone (OR 2.03; 95%CI 1.33 to 3.09) There was no strong evidence for departure from linearity for any of the biomarker-outcome associations (Table 7-3; Supplementary Table C-4complete-case analysis results).

Since we neither observed an association between adiposity measures and TNF-α (Table 7-2) nor between TNF-α and EC risk (Table 7-3), we did not include this biomarker in our mediation analysis.

Table 7-3 Estimated exposure-outcome and mediator-outcome associations; multiple imputation analysis

Body Mass Index	**≥30 vs <25 (analysis excluded women categorized as overweight)**			
OR (95% CI)	3.00	(1.70 to	5.28)	
Waist circumference	**>88 vs ≤80 (analysis excluded women categorized as overweight)**			
OR (95% CI)	2.08	(1.29 to	3.35)	
Waist-hip ratio	**>0.84 vs ≤0.78 (analysis excluded women categorized as overweight)**			
OR (95% CI)	1.57	(0.93 to	2.64)	
Biomarker	**Per doubling concentration**			**P value for evidence against linearity**
Adiponectin				
Model 1 (adj for confounders + BMI)	0.65	(0.47 to	0.90)	0.05
Interleukin-6				
Model 1 (adj for confounders + BMI)	1.11	(0.87 to	1.43)	
Model 2 (adj for confounders + BMI + adiponectin)	1.05	(0.82 to	1.36)	0.36
Interleukin-1 receptor antagonist				
Model 1 (adj for confounders + BMI)	1.15	(1.02 to	1.31)	
Model 2 (adj for confounders + BMI + adiponectin)	1.14	(1.00 to	1.29)	0.57
Tumour necrosis factor-α				
Model 1 (adj for confounders + BMI)	0.99	(0.80 to	1.23)	
Model 2 (adj for confounders + BMI + adiponectin)	0.97	(0.78 to	1.21)	0.83
Tumour necrosis factor-receptor 1				
Model 1 (adj for confounders + BMI)	1.05	(0.52 to	2.12)	
Model 2 (adj for confounders + BMI + adiponectin)	1.05	(0.51 to	2.16)	0.30
Tumour necrosis factor-receptor 2				
Model 1 (adj for confounders + BMI)	0.97	(0.55 to	1.70)	
Model 2 (adj for confounders + BMI + adiponectin)	0.97	(0.55 to	1.72)	0.24
C-reactive protein				
Model 1 (adj for confounders + BMI)	1.09	(0.92 to	1.29)	
Model 2 (adj for confounders + BMI + adiponectin)	1.06	(0.89 to	1.25)	0.54
C-peptide				
Model 1 (adj for confounders + BMI)	1.16	(0.84 to	1.59)	
Model 2 (adj for confounders + BMI + adiponectin + IL1-RA)	1.00	(0.72 to	1.40)	0.77
Free estradiol				
Model 1 (adj for confounders + BMI)	1.55	(1.08 to	2.24)	
Model 2 (adj for confounders + BMI + adiponectin + IL1-RA + C-peptide)	1.38	(0.94 to	2.02)	0.11
Estrone				
Model 1 (adj for confounders + BMI)	2.13	(1.41 to	3.21)	
Model 2 (adj for confounders + BMI + adiponectin + IL1-RA + C-peptide)	2.03	(1.33 to	3.09)	0.57

Models were adjusted for age at blood collection, time at blood collection, fasting status at blood draw, recruitment centre, age at menarche, number of full-term pregnancies, history of oral contraceptive use, history of hormone therapy, physical activity, and smoking status. Models for biomarkers were additionally adjusted for body mass index continuous.

Mediation analysis with multiple mediators: The OR^{NIE} for the BMI (≥30*vs.* ≥18.5-<25 kg/m^2)-EC association through all biomarkers was 1.95 (95%CI 1.01–3.74), corresponding to ~72% of the association mediated by all biomarkers. This NIE was further decomposed into an NIE through pathways originating with reduced adiponectin

and increased inflammation biomarkers (OR^{NIE} 1.53; 95%CI 0.89–2.62; 46%PM); through C-peptide, beyond [the potential influences] of adiponectin and inflammation biomarkers (OR^{NIE} 1.05; 95%CI 0.88–1.24; 5%PM); and through estrogens (OR^{NIE} 1.22; 95%CI 0.89–1.67; 21%PM) beyond adiponectin, inflammation biomarkers, and C-peptide. The NIE through adiponectin and inflammation biomarkers was further decomposed into an NIE through pathways originating with adiponectin (OR^{NIE} 1.35; 95%CI 1.06–1.73; 33%PM) and through inflammation biomarkers beyond adiponectin (OR^{NIE} 1.13; 95%CI 0.71–1.80; 13%PM). The estimated OR^{NDE} not through any of the biomarkers was 1.29 (95%CI 0.54–3.09). The OR^{NIE} point estimates for inflammation biomarkers, C-peptide, and estrogens were suggestive of moderate to weak increase in EC OR, but the 95%CIs were wide and included a decreased risk (Table 7-4; Supplementary Table C-5 complete-case analysis results).

For waist circumference (>88*vs.*≤80 cm), the ORs^{NIE} were 1.73 (95%CI 1.04–2.90; 76%PM) through all biomarkers; 1.56 (95%CI 1.01–2.42; 61%PM) through reduced adiponectin and increased inflammation biomarkers; 1.32 (95%CI 1.03–1.68; 38%PM) through reduced adiponectin; 1.19 (95%CI 0.83–1.69; 24%PM) through inflammation biomarkers beyond adiponectin; 1.03 (95%CI 0.89–1.19; 4%PM) through C-peptide beyond adiponectin and inflammation biomarkers; and 1.08 (95%CI 0.88–1.33; 10%PM) through estrogens beyond adiponectin, inflammation biomarkers, and C-peptide. The OR^{NDE} not through any of the biomarkers was 1.19 (95%CI 0.59–2.41) (Table 7-4, Supplementary Table C-5 complete-case analysis). The 95% CIs around the OR^{NIE} point estimates for inflammatory markers, C-peptide, and estrogens were wide.

Mediation patterns were similar for continuous exposures (Supplementary Table C-6) and after excluding cases diagnosed within two years post-recruitment (Supplementary Table C-7).

Table 7-4 Estimated natural direct and indirect effects using sequential mediation analysis; analyses excluded women categorized as overweight; multiple imputation analysis

	Body Mass Index ≥30 vs. <25 kg/m2				Waist Circumference >88 vs. ≤80 cm			
Number of cases/ number of controls	102/185				131/212			
	Odds Ratio (95% CI)			% mediated on log odds scale	Odds Ratio (95% CI)			% mediated on log odds scale
Total effect (estimated as the product of natural direct and indirect effects)	2.51	(1.26 to	5.02)		2.07	(1.20 to	3.55)	
Natural indirect effect through all the biomarkers	1.95	(1.01 to	3.74)	72%	1.73	(1.04 to	2.90)	76%
Natural indirect effect through reduced adiponectin and increased inflammation	1.53	(0.89 to	2.62)	46%	1.56	(1.01 to	2.42)	61%
Natural indirect effect through reduced adiponectin levels	1.35	(1.06 to	1.73)	33%	1.32	(1.03 to	1.68)	38%
Natural indirect effect through increased inflammation, beyond the potential influence of adiponectin	1.13	(0.71 to	1.80)	13%	1.19	(0.83 to	1.69)	24%
Natural indirect effect through increased c-peptide levels, beyond the potential influences of adiponectin and inflammation	1.05	(0.88 to	1.24)	5%	1.03	(0.89 to	1.19)	4%
Natural indirect effect through increased free estradiol and estrone levels, beyond the potential influences of adiponectin, inflammation, and c-peptide	1.22	(0.89 to	1.67)	21%	1.08	(0.88 to	1.33)	10%
Natural direct effect not through any of the biomarkers	1.29	(0.54 to	3.09)		1.19	(0.59 to	2.41)	

Confounders included age at blood collection, country region, age at menarche, number of full-term pregnancies, history of oral contraceptive use, history of hormone therapy, physical activity, and smoking status. Mediation analyses were performed within each of the 20 imputed datasets and resulting estimates and standard errors (derived from 1,000 bootstrap draws) were pooled to obtain the multiple imputation estimates and 95% confidence intervals.

Discussion

In this study of postmenopausal women, reduced adiponectin and increased inflammation biomarkers, C-peptide, and estrogens mediated most (>70%) of the effect of adiposity on EC risk. In the sequential mediation analysis, the largest mediating effect was observed for pathways originating with adiponectin. Based on the point estimates, depending on the measure of adiposity used, the second most important pathway was either inflammation (waist circumference) or estrogens (BMI).

We believe this is the first study to investigate the mediating role of multiple biomarkers in the adiposity-EC association using formal mediation analysis. We had measures for a range of biomarkers representing the pathways of interest. However, because this was secondary analysis of an existing study, we were limited to biomarkers that had been measured and these may not have fully captured the entire physiological impact of these pathways. For the insulin pathway, we had measures for IGF binding protein (IGFBP)-1 and IGFBP-2 but did not include them because they do not have a clear relationship with adiposity or EC (173). Additionally, the observed mediating effect through biomarkers might have been influenced by the quality of the measurements (the intra-batch coefficient of variation (CV) ranged from 2.6% (C-peptide) to 15% (IL-1Ra, TNF-α), and the inter-batch CV from <8% (adiponectin, TNF-R1, TNF-R2) to 27.7% (IL-1IRA) (77, 78, 94, 113, 140), and temporal stability of biomarkers.

The sequential mediation analysis circumvented the assumption of no exposure-induced mediator-outcome confounding and permitted quantifying the indirect effect through all the biomarkers, and decomposing this into path-specific indirect effects without assuming the biomarkers were independent (129). The sequential mediation analysis

relied on an assumed causal ordering of the pathways, which was based on existing evidence (8, 54, 149).

We controlled for the known confounders of the associations between exposure, mediators, and outcome, but residual confounding cannot be ruled out. We were also limited in exploring the potential influence of unmeasured confounding on our results, because sensitivity analysis and bias correction approaches for settings with multiple mediators are not developed yet. Although our analysis assumed that it was adiposity that influenced biomarker levels, we cannot be certain about the temporal exposure-mediator ordering, because they were measured cross-sectionally. Preclinical cancers might have also influenced measures of adiposity and biomarkers. However, results from a sensitivity analysis that excluded cases diagnosed within two years post-recruitments were comparable to our primary analysis. We used a regression-based mediation analysis that could be adapted to the matched design (128). A limitation was that it did not allow including possible exposure-mediator and mediator-mediator interactions (129).

The predominant hypothesis in explaining the adiposity-EC association in postmenopausal women is increased estrogen production in adipose tissue (112, 190). To explore the mediating role of estrogens, two studies assessed the degree to which adjusting for estradiol attenuated the BMI-EC association (113, 114). The adjustment weakened but did not entirely explain the association (OR for BMI ≥30*vs.* <25kg/m^2 from 3.97 (95%CI 2.54-6.21) to 2.25 (95%CI 1.33-3.81) (114), and from 2.67 (95%CI 1.63-4.37) to 2.09 (95%CI 1.22-3.57) (113)), suggesting that other pathways are also likely at play. Similarly, we observed weak evidence of a small to moderate mediating effect through estrogens beyond the potential influence of preceding pathways.

Hyperinsulinemia might mediate the adiposity-EC association (182). However, we found almost no mediating effect through C-peptide beyond adiponectin and inflammation. Triglyceride glucose product (TyG) was used as a proxy for insulin resistance in a pooled study of six European cohorts. A negligible proportion (3.6%) of the total effect of BMI (≥30*vs.* ≥18.5-<25 kg/m^2; hazard ratio (HR) 2.61; 95%CI 2.29-2.99) was mediated by TyG (HR^{NIE} 1.04; 95%CI 0.99-1.09) (134). In both studies, there was little association between TyG or C-peptide and EC after adjusting for BMI. However, other studies have found associations between fasting insulin (49, 90) and C-peptide (93)) and EC, after adjustment for BMI. Put together with the likely limitations of fasting and non-fasting C-peptide in capturing insulin resistance (90), such discrepancies warrant additional research to elucidate the mediating effect of the insulin resistance pathway.

In our analysis, the largest mediating effect was observed through reduced adiponectin and increased inflammation biomarkers. A further decomposition of this indirect effect suggested that a larger mediating effect was through adiponectin, with indication of a smaller mediating effect through inflammation beyond the potential influence of adiponectin. Existing evidence for inverse adiposity-adiponectin (191) and adiponectin-EC associations (192) support our observation. Based on a principal-component factor analysis, data from the original case-control study previously demonstrated a reduction in the OR for the association between BMI (≥30*vs.*<25 kg/m^2) and postmenopausal EC from 2.73 (95%CI, 1.66-4.50) to 1.65 (95%CI, 0.92-2.98) (~50% reduction on log(OR) scale) after adjusting for a factor with >|40%| loading for adiponectin, together with CRP, C-peptide, IGFBP-1, IGFBP-2, SHBG, and HDL cholesterol (118).

In summary, about 70% of the increased odds of EC risk in women with obesity compared with normal weight was mediated together through adiponectin, inflammation, C-peptide, and estrogens. We applied a novel mediation analysis approach to quantify the mediating effects through these pathways jointly, and the path-specific indirect effects. The applied method was able to handle multiple correlated biomarkers without assuming independence. Pathways originating with reduced adiponectin had the most important role in explaining the link. Future studies, with larger sample sizes and a range of biomarkers reflecting the pathways, are needed to replicate findings from this study.

Disclaimer: Where authors are identified as personnel of the International Agency for Research on Cancer / World Health Organization, the authors alone are responsible for the views expressed in this article and they do not necessarily represent the decisions, policy or views of the International Agency for Research on Cancer / World Health Organization.

Funding: This work was supported by the World Cancer Research Fund (grants 2003/18 and 2007/13); Australian National Health and Medical Research Council grants (1150591 to SGD and 1104975 to JAS); and Melbourne Research Scholarship (to SGD). MMB is the recipient of an Australian Research Council Discovery Early Career Award (project number DE190101326) funded by the Australian Government. The work reported in this paper was undertaken by SGD while hosted by the International Agency for Research on Cancer (IARC) and funded partially by a PhD Scholarship by IARC. The coordination of EPIC is financially supported by the European Commission and the International Agency for Research on Cancer. The national cohorts are supported by the following: Danish Cancer Society (Denmark); Ligue Contre le Cancer, Institut Gustave Roussy, Mutuelle Générale de l'Education Nationale, and Institut National de la Santé et de la Recherche Médicale (France); Deutsche Krebshilfe, Deutsches Krebsforschungszentrum, and Federal Ministry of Education and Research (Germany); Stavros Niarchos Foundation and Hellenic Health Foundation (Greece); Italian Association for Research on Cancer and National Research Council (Italy); Dutch Ministry of Public Health, Welfare and Sport, Netherlands Cancer Registry, LK Research Funds, Dutch Prevention Funds, Dutch ZON (Zorg Onderzoek Nederland), World Cancer Research Fund, and Statistics Netherlands (the Netherlands); European Research Council and Nordforsk, Nordic Centre of Excellence Programme on Food, Nutrition and Health (Norway); Health Research Fund, Regional Governments of Andalucía, Asturias, Basque Country, Murcia, and Navarra, and Instituto de Salud Carlos III Red Temática de Investigación Cooperativa en Salud (Spain); Swedish Cancer Society, Swedish Scientific Council, and Regional Governments of Skåne and Västerbotten (Sweden); and Cancer Research UK, Medical Research Council, Stroke Association, British Heart Foundation, Department of Health, Food Standards Agency, and Wellcome Trust (United Kingdom).

Acknowledgements: We thank the National Institute for Public Health and the Environment (RIVM), Bilthoven, the Netherlands for their contribution and ongoing support to the EPIC Study. This study also made use of data from the EPIC-Norfolk cohort. We thank Britta Lederer, Sigrid Henke, Nivonirina Robinot, Priscilia Amouyal, Josiane Bouzac, Thomas Cler, and David Achaintre for their meticulous work in performing the immunoassays used for the various studies included in this reanalysis.

Conflict of interest: The authors declare no potential conflicts of interest.

Chapter 8: Adiposity and colorectal, estrogen receptor-positive breast, and endometrial cancer risk in postmenopausal women in the WHI-OS

In this chapter, an analysis of a case-cohort study within the Women's Health Initiative Observational Study (WHI-OS) has been carried out to investigate the mediating roles of adiponectin, leptin, resistin, hepatocyte growth factor (HGF), plasminogen activator inhibitor (PAI), IL-6, and CRP (to reflect the inflammatory status), fasting insulin (to reflect hyperinsulinemia), and total estradiol (to reflect alterations in sex-steroid hormone levels) in the effect of adiposity on ER-positive breast, endometrial, and colorectal cancer risk in postmenopausal women. This chapter includes the draft of an in-progress manuscript. The work presented here was conducted after the proposal for the analysis was approved by the WHI Publications and Presentations (P&P) committee (proposal number 3809). The proposal authors are S. Ghazaleh Dashti, Marc J. Gunter, Gloria Y.F. Ho, Howard D. Strickler, and Sylvia Wassertheil-Smoller. Professor Sylvia Wassertheil-Smoller is also the sponsoring WHI Principal Investigator. At the time of submission of this thesis, S Ghazaleh Dashti, Julie A Simpson, Vivian Viallon, Amalia Karahalios, Margarita Moreno-Betancur, Dallas R English, and Marc J Gunter have read and contributed to the manuscript. However, this is not the final author list and it will be expanded prior to submission of the manuscript to a nominated journal. Maintaining the confidentiality of the findings in this paper, is, kindly requested.

For consistency, spelling, referencing, tables, and figure numbers have been modified. Supplementary material related to this publication are presented in Appendix D.

Adiposity and breast, endometrial, and colorectal cancer risk in postmenopausal women: quantification of the mediating effects of inflammation, fasting insulin, and estradiol

S Ghazaleh Dashti[1,2], Julie A Simpson[1], Vivian Viallon[2], Amalia Karahalios[1,3], Margarita Moreno-Betancur[4,5], Dallas R English[1,6], Marc J Gunter[2]

1 Centre for Epidemiology and Biostatistics, Melbourne School of Population and Global Health, The University of Melbourne, Melbourne, Victoria, Australia
2 Section of Nutrition and Metabolism, International Agency for Research on Cancer (IARC), Lyon, France
3 School of Public Health and Preventive Medicine, Monash University, Melbourne, Victoria, Australia
4 Clinical Epidemiology and Biostatistics Unit, Murdoch Children's Research Institute, Melbourne, Australia
5 Department of Paediatrics, The University of Melbourne, Melbourne, Victoria, Australia
6 Cancer Epidemiology and Intelligence Division, Cancer Council Victoria, Melbourne, Victoria, Australia

Corresponding author:
S Ghazaleh Dashti
Centre for Epidemiology and Biostatistics
Melbourne School of Population and Global Health
Level 3, 207 Bouverie Street
The University of Melbourne VIC 3010 Australia
Email: seyedeh.dashti@unimelb.edu.au

Abbreviations: ER estrogen receptor; BC breast cancer; EC endometrial cancer; CRC colorectal cancer; IIE interventional indirect effect; IDE interventional direct effect; RR risk ratio; TE total effect; SHBG sex hormone binding globulin; IGF-1 insulin-like growth factor-1; WHI-OS Women's Health Initiative Observational Study; HGF hepatocyte growth factor; PAI plasminogen activator inhibitor-1; TNF-α tumour necrosis factor-α; IL-6 interleukin-6; CRP C-reactive protein; BMI body mass index; WHR waist-hip ratio; GMR geometric mean ratio; CI confidence interval; NSAID nonsteroidal anti-inflammatory drug; HR hazard ratio; RR risk ratio; IQR interquartile range; CV coefficient of variations; ICC intraclass correlation coefficient; MCCS Melbourne Collaborative Cohort Study

Abstract

Background

Mechanisms underlying the effect of adiposity on cancer risk are incompletely understood. We aimed to quantify the mediating roles of biomarkers representing inflammatory status, fasting insulin, and estradiol in the effect of adiposity on estrogen receptor (ER)-positive breast (BC), endometrial (EC), and colorectal (CRC) cancer risk in postmenopausal women.

Methods

We used data from a case-cohort study within the Women's Health Initiative-Observational Study. Interventional indirect (IIE) and direct (IDE) effects were estimated using a regression-standardisation approach.

Results

The study included 187 BC cases, 287 controls; 98 EC cases, 202 controls; and 193 CRC cases, 288 controls. For BC, the adjusted risk ratio (RR) for total effect (TE) for waist circumference (>88vs≤80cm) was 1.68 (95%CI, 0.96–2.79). The RRs^{IIE} were 2.09 (1.15–3.46) through all biomarkers, 1.10 (0.61–1.96) though inflammation, 1.67 (1.08–2.65) through insulin, 1.14 (1.02–1.39) through estradiol, and RR^{IDE} not through biomarkers was 0.80 (0.39–1.71). For EC, RR^{TE} was 2.24 (1.03–5.03), RRs^{IIE} were 2.01 (0.86–3.97) through all biomarkers, 1.40 (0.58–3.06) though inflammation, 1.13 (0.61–1.83) through insulin, 1.27 (1.04–1.69) through estradiol, and RR^{IDE} was 1.11 (0.43–3.75). For CRC, RR^{TE} was 1.61 (0.94–2.58), RRs^{IIE} were 1.61 (0.95–2.52) through all biomarkers, 1.37 (0.82–2.38) though inflammation, 1.16 (0.77–1.69) through fasting insulin, 1.01 (0.89–1.14) through estradiol, and RR^{IDE} was 1.00 (0.51–1.94).

Conclusion

The assessed biomarkers explained the effect of adiposity on the risk of the three cancers to a large extent, but their importance varied between the cancers.

Introduction

Excess body fatness is an established risk factor for several cancers, including estrogen receptor (ER)-positive postmenopausal breast, endometrial, and colorectal cancers (1, 2). Dysregulated adipocytokine production and chronic low-grade inflammation, hyperinsulinemia, and disturbed concentration of bioavailable sex-steroid hormones are three of the major pathways hypothesized to underlie this effect (1).

Chronic low-grade inflammation is a known consequence of excess adiposity, in which the production of pro-inflammatory adipocytokines and cytokines increases and the production of adiponectin (an adipocytokine with anti-inflammatory properties) decreases (1). This inflammatory status can promote carcinogenesis by enabling cell proliferation, apoptosis, and angiogenesis (5), impairing insulin signalling thereby inducing hyperinsulinemia (9), and disturbing bioavailable sex-steroid hormone levels through enhancing aromatase expression (54) and downregulating sex hormone binding globulin (SHBG) production (12). Individuals with obesity are also more likely to have insulin resistance and therefore, chronic hyperinsulinemia (8). Insulin has mitogenic and anti-apoptotic effects (1, 13, 14). Hyperinsulinemia might also affect cancer risk through increasing bioavailable insulin-like growth factor-1 (IGF-1) levels, which has growth promoting effects (14, 15), and increasing bioavailable estrogens through enhancing aromatase expression (10) and possibly downregulating SHBG production (12). Estrogens have proliferative effects on breast and endometrial tissues (1), while there is some evidence that they might have antiproliferative effects on colon tissue through their interaction with ERβ expressed in the colonic epithelium (1).

The existing case-cohort study of the inflammatory biomarkers and breast, endometrial, and colorectal cancers in postmenopausal women within the Women's Health Initiative-

Observation Study (WHI-OS) (68, 76, 193) is one of the few studies we are aware of that has measurements available for a range of adipocytokines and cytokines (including adiponectin, leptin, resistin, hepatocyte growth factor (HGF), plasminogen activator inhibitor-1 (PAI-1), tumour necrosis factor-α (TNF-α), interleukin-6 (IL-6), C-reactive protein (CRP)) as well as for fasting insulin and estradiol (68, 76, 193). This study is a unique resource for further exploring the role of inflammatory status (as captured by the measured adipocytokines and cytokines), hyperinsulinemia, and estradiol on the effect of adiposity on ER-positive breast, endometrial, and colorectal cancers in postmenopausal women. Firstly, it allows simultaneous assessment of the importance of these three potentially mechanistic pathways, which as described above are interrelated. Secondly, it allows comparing the possibly different mediating role of biomarkers reflecting these pathways between the three cancer types (13). In the present study, we used these data and a multiple-mediator causal mediation analysis approach (130) to quantify the mediating effects of these three pathways on the effect of adiposity on each cancer site, while taking the dependence between the mediators into account.

Methods

The Women's Health Initiative Observational Study (WHI-OS) is a study of 93,676 postmenopausal women aged 50 to 79 years old, recruited from 40 centres across the United States from 1993 to 1998 (141). At recruitment (baseline), participants provided written informed consent and completed a questionnaire on demographic and lifestyle factors, medical history, and history of medication use. Height, weight, and waist and hip circumference were measured during a physical examination, when fasting blood samples were also collected, processed, and plasma and serum stored at -80°C (141).

The present study made use of an existing case-cohort study on inflammation, fasting insulin, and estradiol and the risk of endometrial, breast, and colorectal cancers (59, 68,

76). Cancer cases were ascertained based on annual self-administered questionnaires and confirmed via centralized review of medical reports. For the case-cohort study, follow-up ended on 29 February 2004, as of when 1.6% of all participants were lost to follow-up and 4.7% had died.

Participants

Women were included in the original case-cohort study if they had more than one year of follow-up, were not diagnosed with endometrial, breast, or colorectal cancer in the first year, and were not taking diabetes medication at baseline (59, 68, 76). Also, for endometrial cancer, eligible women had no history of hysterectomy and were not taking hormone therapy at baseline (76). For the present study, women were additionally ineligible if at baseline they had history of any cancer diagnosis (except keratinocyte skin cancer), had unknown diabetes status or history of diagnosis with diabetes, had unknown hormone therapy use or currently used hormone therapy, or had BMI <18.5 kg/m^2. We also excluded participants with missing values for baseline body size measurements (BMI, waist circumference, waist-hip ratio), selected confounders (see *Confounder Selection*), or biomarkers (see *Biomarkers*) (Table 8-1).

All estrogen receptor (ER)-positive breast, endometrial, and colorectal cancer cases diagnosed between one year after recruitment and 29 February 2004 and a random sample of all eligible women in the cohort regardless of their subsequent cancer diagnosis status (sub-cohort) were included. For the mediation analysis (see *Mediation Analysis*), because methods for mediation analysis with multiple mediators and time-to-event outcomes are at their inception (144), we treated the data as if the design followed an unmatched case-control study with cumulative sampling. For each cancer site, the control group consisted of women in the sub-cohort who had not been diagnosed with the cancer of interest

(endometrial, breast, or colorectal cancer), and had not died or been lost to follow-up before the end of follow-up (Table 8-1).

Table 8-1 Number of women affected and unaffected with endometrial, ER-positive breast, and colorectal cancer in the original case-cohort study, number to be excluded, and the final number in the analytic dataset for this study

		ER positive Breast	Endometrial Cancer	Colorectal Cancer
Case-cohort study	Number affected/ Number in the **sub-cohort**	552/821 (13 with ER-positive breast cancer)	151/300 (2 with endometrial cancer)	457/839 (7 with colorectal cancer)
	Did not meet eligibility criteria			
	Had history of any cancer diagnosis (except keratinocyte skin cancer)	43/128	35/47	78/129
	Unknown diabetes status at baseline	0/0	0/0	1/1
	Diagnosed with diabetes at baseline	8/7	3/2	8/7
	Unknown hormone therapy intake at baseline	0/0	0/0	0/0
	Used hormone therapy at baseline	278/333	0/0	145/333
	With BMI<18.5 kg/m^2	1/4	1/2	2/4
	Excluded due to missing data			
	missing baseline body size measurements	3/15	1/10	5/15
	missing confounder data	17/15	2/12	7/16
	missing any biomarker data	15/20	11/13	18/21
	Number affected/ Number in the **sub-cohort** in the analytic dataset	**187/311** (4 with ER-positive breast cancer)	**98/214** (1 with endometrial cancer)	**193/ 313** (3 with colorectal cancer)
Case-control study with cumulative sampling for mediation analyses	Number affected/ Number **not affected with cancer**	187/307	98/213	193/310
	Excluded controls who died or were lost to follow-up before the end of the follow-up time of the case-cohort study	0/20	0/11	0/22
	Number affected/ Number **not affected with cancer** in the analytic dataset for mediation analysis	**187/287**	**98/202**	**193/288**

Abbreviations ER estrogen receptor

Biomarkers

Biomarkers of inflammatory status were measured in baseline plasma samples and included adiponectin, leptin, resistin, plasminogen activator inhibitor-1 (PAI-1), hepatocyte growth factor (HGF), tumour necrosis factor-α (TNF-α), interleukin-6 (IL-6) and C-reactive protein (CRP) (59, 68, 76). Fasting insulin and estradiol were measured in serum (83, 90, 119). Assay methods for measuring the biomarkers have been described previously (59, 68, 76, 83, 90, 119).

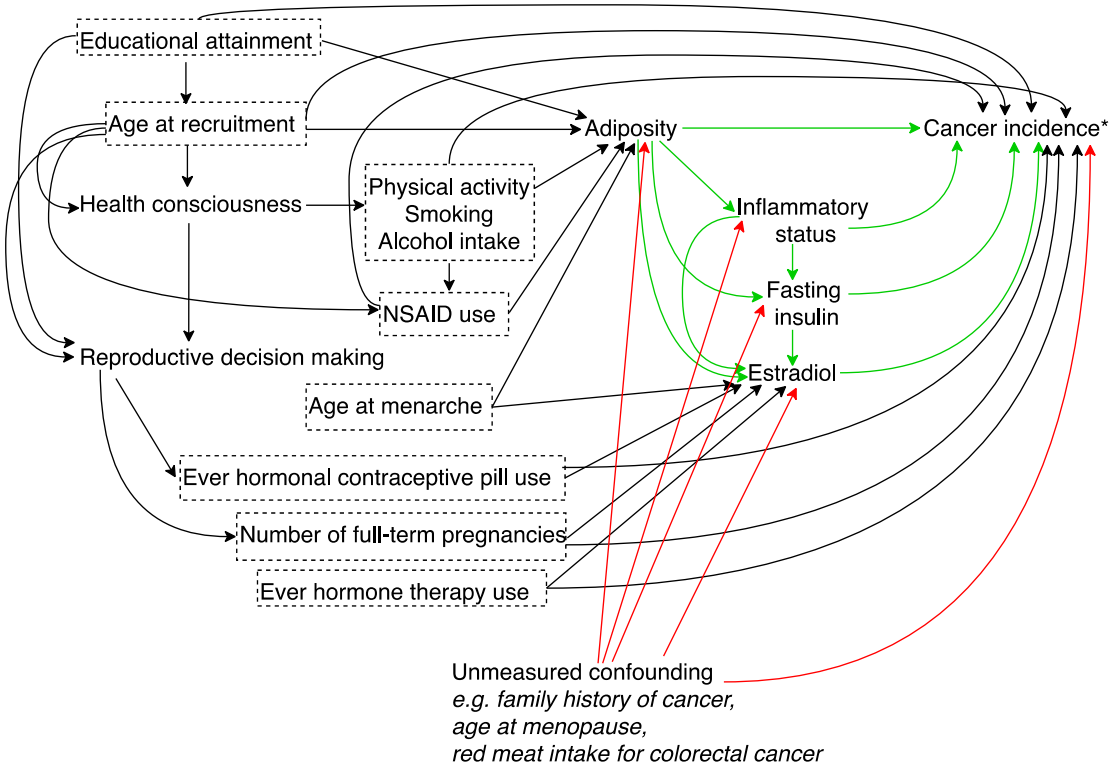
Statistical analysis

Separate analyses for each measure of the adiposity exposure (waist circumference, BMI, and waist-to-hip ratio) and each outcome (diagnosis with ER-positive breast, endometrial, and colorectal cancer) were performed. In separate sets of analyses, diagnosis with ER-positive breast, endometrial, and colorectal cancers were treated as the outcomes. Waist circumference, BMI, and waist-to-hip ratio were used as measures of adiposity (exposure) in separate analyses. To help with the interpretation of the estimated mediated effects (see *Mediation Analysis*), adiposity measures were treated as binary comparing women categorized as with obesity vs normal weight (BMI ≥30 vs <25; waist circumference >88 vs ≤80 cm; waist-hip ratio third (>0.83) vs first (≤0.77) tertile based on the distribution of waist-hip ratio in the sub-cohort (25)).

Confounder selection: The following potential confounders (common causes of exposure-outcome, mediator-outcome, or exposure-mediator associations (155)) were identified *a priori* based on existing evidence (34, 41, 45, 194) and included: age at baseline, educational attainment, alcohol intake, smoking status, physical activity, history of nonsteroidal anti-inflammatory drugs use, age at menarche, parity, ever use of

hormonal contraceptives, ever use of hormone therapy, age at menopause, red meat intake (only for colorectal cancer), and family history of cancer (Figure 8-1). Information on all confounders were collected at baseline. Final analyses did not include the potential confounders with more than 2% missing values (age at menopause, red meat intake, family history of cancer) which was decided on the basis that all had weak correlations with adiposity measures and biomarkers (Supplementary Table D-1).

Figure 8-1 Assumed causal diagram for quantifying the mediating effects of inflammatory status, fasting insulin, and estradiol on adiposity-cancer (estrogen receptor positive breast, endometrial, and colorectal cancer) association in postmenopausal women



Abbreviations NSAID Nonsteroidal anti-inflammatory drug
*To avoid redundancy, a single causal diagram has been presented for the three cancers included in this study (i.e. estrogen-receptor positive breast, endometrial, and colorectal cancer). Also, to simplify the diagram, we only included the arrows that were sufficient to flag a variable as a common cause (i.e. confounder) of exposure-outcome, or exposure-mediator, or mediator-outcome. Therefore, this is not a causal diagram *per se*. It is assumed that all potential confounders preceded the exposure, mediators, and the outcome. The diagram was developed with reference to the literature (34, 41, 45, 194) and expert opinion. The green arrows depict the direct and indirect pathways we were interested in, the red arrows the pathways that might have introduced biased due to unmeasured confounding, and the black arrows the backdoor pathways that were blocked by conditioning on the boxed variables. Dashed boxes indicate variables that were included in multivariable analyses.

Adiposity-biomarker associations: For each biomarker, the geometric mean ratio (GMR) and 95% confidence intervals (CI) were estimated in relation to adiposity measures (comparing obesity vs normal weight) using linear regression models fitted to the log-transformed biomarker variable and adjusted for potential confounders (Figure 8-1). The models were fitted for all women selected into the sub-cohort who met the eligibility criteria for this study.

Adiposity-cancer and biomarker-cancer associations: Cox regression models with follow-up time as the time axis were used to estimate the hazard ratio (HR; 95%CI) for the association between adiposity or biomarkers and each cancer site. The case-cohort design was taken into account using the Prentice approach with robust standard errors (170). For each cancer site, observation time for the sub-cohort participants started at recruitment and ended at the time of diagnosis of the cancer of interest, death, lost to follow-up, or end of follow-up, whichever happened first. All models were conditional on the baseline confounders. Models for biomarkers also included waist circumference, as the measure for adiposity, which might have confounded the biomarker-cancer association. Additionally, models for fasting insulin also included CRP (as a proxy for inflammatory status) to take the potential confounding effect of this biomarkers on fasting insulin-cancer association into account. Similarly, models for estradiol additionally included CRP and fasting insulin as potential confounders (Figure 8-1).

The linearity of the mediator-cancer associations was investigated by fitting the biomarker variables as restricted cubic splines (with 2 degrees of freedom).

For each cancer site, we only performed mediation analyses (see *Mediation analysis* below) for the adiposity measures for which there was evidence for a positive association based on the point estimate and 95%CIs.

Mediation analysis: As previously explained, for mediation analysis the study design was treated as an unmatched case-control study with cumulative sampling. To assess that analysing the data as such was not likely to introduce material bias, risk ratios (RR) and 95%CIs estimated from the logistic regression models for the adiposity-cancer and biomarker-cancer associations were compared with the HRs (and 95%CIs) from Cox models. Risk ratios were estimated from the logistic regression models based on the formula $risk = \frac{1}{(1+\exp(-\log(\text{odds})))}$ (156). Prior to calculating risks, the $\log(\frac{1}{sampling\ fraction\ of\ controls})$ was subtracted from the constant term to take the case-cohort design into account (174).

We performed a causal mediation analysis (130) to estimate the interventional indirect (IIE) and direct effects (IDE). For a given biomarker (or set of biomarkers) M, IIE is defined as the average relative (or absolute) change in the probability of cancer incidence for women with obesity, while intervening to shift the distribution of M from its observed distribution in women with obesity given baseline confounders to its observed distribution in women with normal weight given baseline confounders. The IDE is defined as the average relative (or absolute) change in the probability of cancer incidence for women with obesity versus normal weight, when fixing the distribution of M to its observed distribution in women with normal weight given baseline confounders (130, 132). The following IIEs were estimated: 1) through all biomarkers; 2) through inflammatory status (captured by reduced levels of adiponectin and increased levels of pro-inflammatory adipocytokines and cytokines); 3) through fasting insulin; and 4) through estradiol; and an IDE to capture the effect of adiposity on cancer risk not through any of the included biomarkers. We also estimated the difference between the total effect

of adiposity on cancer (TE) and the product (for risk ratios) or sum (for risk differences) of the estimated IDE and IIEs, which would capture an IIE through the dependence and interaction between the mediators (130). More detailed interpretations of the IIEs and IDE are summarized in Table 8-2 (130, 132). This mediation analysis method is able to handle multiple correlated biomarkers without assuming that they are independent.

The IDE and IIEs were estimated on the risk ratio and risk difference scales based on a series of linear regression models for the mediators, , conditional on the adiposity measure and confounders (fitted to the sub-cohort to take the case-cohort design into account (155)), and a logistic regression model for cancer incidence conditional on the adiposity measure, all biomarkers, and confounders. Due to limited sample size, exposure-mediator or mediator-mediator interactions were not included in any of the models. Additional detail on these models and how they were combined to estimate the IIE and IDE are provided in Supplementary Table D-2. The estimation of the effects followed a simulation-based, regression-standardisation approach where the estimates were averaged based on a large number of Monte-Carlo draws (~1000 times the sample size as proposed by Vansteelandt and Daniel (130)) from the mediator distribution, with the distributions of the confounders set to their observed distributions in the sample (130). Standard errors (and 95%CIs) were estimated using 2,000 bootstrap samples.

All analyses were complete case and performed in Stata version 15 (158).

Table 8-2 Interpretations of the interventional indirect and direct effects

#	Effect	Interpretation
1	The interventional **indirect** effect through **all the biomarkers**	Average relative (or absolute) change in the probability of cancer incidence for women with obesity, while intervening to shift the joint distribution of all biomarkers from their observed joint distribution in women with obesity adjusted for baseline confounders to their observed joint distribution in women with normal weight adjusted for baseline confounders
2	Interventional **indirect** effect through inflammatory status	Average relative (or absolute) change in the probability of cancer incidence for women with obesity, while intervening to shift the joint distribution of all biomarkers of inflammatory status from their observed joint distribution in women with obesity adjusted for baseline confounders to their observed distribution in women with normal weight adjusted for baseline confounders and fixing the distribution of fasting insulin and estradiol to their observed distributions in women with normal weight adjusted for baseline confounders
3	Interventional indirect effect through fasting insulin	Average relative (or absolute) change in the probability of cancer incidence for women with obesity, while intervening to shift the distribution of fasting insulin from its observed distribution in women with obesity adjusted for baseline confounders to its observed distribution in women with normal weight adjusted for baseline confounders and fixing the joint distribution of biomarkers of inflammatory status to their observed joint distribution in women with obesity adjusted for baseline confounders and the distribution of estradiol to its observed distribution in women with normal weight adjusted for baseline confounders
4	Interventional indirect effect through estradiol	Average relative (or absolute) change in the probability of cancer incidence for women with obesity, while intervening to shift the distribution of estradiol from its observed distribution in women with obesity adjusted for baseline confounders to its observed distribution in women with normal weight adjusted for baseline confounders and fixing the joint distribution of biomarkers of inflammatory status and fasting insulin to their observed distributions in women with obesity adjusted for baseline confounders
5	Interventional direct effect not through any of the biomarkers	Average relative (or absolute) change in the probability of cancer incidence for women with obesity versus normal weight, when fixing the joint distribution of all biomarkers to their observed joint distribution in women with normal weight adjusted for baseline confounders
6	Interventional indirect effect through dependence of biomarkers	The difference (relative or absolute) between the total effect and the product (or sum) of direct and indirect effects (effects 1 and 5 or euqally effects 2 through to 5). This effect captures the average relative (or absolute) change in the probability of cancer incidence for women with obesity through the dependence and interaction between all biomarkers. This effect is generally expected to diverge from the null in the presence of mediator-mediator and exposure-mediator interactions, if the models have included these interaction terms.

Additional analyses: We further estimated the IIEs through each of the adipokines (i.e. adiponectin, leptin, resistin, PAI-1, and HGF), and through the rest of the inflammatory biomarkers jointly.

Results

The datasets for the mediation analyses included 187 ER-positive breast cancer cases and 287 controls (median follow-up 5.9 years (interquartile range 4.2–7.0)), 98 endometrial cancer cases and 202 controls (median follow-up 6.1 years (interquartile range 4.8–7.5)), and 193 colorectal cancer cases and 288 controls (median follow-up 6.0 years (interquartile range 4.4–7.1)) (Table 8-1). At baseline, compared with their controls, women with ER-positive breast cancer had a smaller number of term pregnancies and a smaller proportion of them were never or former drinkers. A slightly larger proportion of women with endometrial cancer had ever used hormone therapy at baseline compared with controls. For colorectal cancer, a smaller proportion of cases were never or former drinkers or had ever used hormone therapy at baseline compared with controls (Table 8-3). For all cancer sites, cases were more likely to have BMI≥30 kg/m^2, waist circumference>88 cm, or waist-hip ratio>0.83 relative to controls (Table 8-3).

Table 8-3 Baseline characteristics of women included in the analytic dataset for mediation analysis

	ER-positive Breast Cancer		Endometrial Cancer		Colorectal Cancer	
	Affected	Unaffected	Affected	Unaffected	Affected	Unaffected
	N=187	N=287	N=98	N=202	N=193	N=288
Age at baseline - years, median (IQR)	66.0 (60.0-70.0)	64.0 (58.0-69.0)	64.5 (60.0-71.0)	63.0 (57.0-69.0)	67.0 (61.0-72.0)	64.0 (58.0-69.5)
Educational attainment						
some college or less	98 (52)	164 (57)	51 (52)	111 (55)	120 (62)	165 (57)
college and above	89 (48)	123 (43)	47 (48)	91 (45)	73 (38)	123 (43)
Alcohol intake						
Never or former drinker	39 (21)	94 (33)	30 (31)	62 (31)	54 (28)	96 (33)
<1 drink per week	52 (28)	93 (32)	33 (34)	70 (35)	70 (36)	91 (32)
1 or more drinks per week	96 (51)	100 (35)	35 (36)	70 (35)	69 (36)	101 (35)
Ever smoked	90 (48)	134 (47)	52 (53)	101 (50)	95 (49)	136 (47)
Physical activity - total METs per week	7.5 (2.0-17.5)	9.2 (2.5-19.8)	11.0 (3.4-18.8)	9.8 (3.0-20.2)	8.5 (2.5-16.3)	8.8 (2.5-19.5)
Used NSAIDs 2 weeks or more	63 (34)	90 (31)	30 (31)	62 (31)	56 (29)	92 (32)
Age at menarche - years	13.0 (12.0-13.0)	13.0 (12.0-13.0)	13.0 (11.0-13.0)	13.0 (12.0-13.0)	13.0 (12.0-13.0)	13.0 (12.0-13.0)
Number of term pregnancies	2.0 (1.0-4.0)	3.0 (2.0-4.0)	3.0 (1.0-3.0)	3.0 (2.0-4.0)	3.0 (1.0-4.0)	3.0 (2.0-4.0)
Ever used oral contraceptives	57 (30)	104 (36)	37 (38)	77 (38)	50 (26)	104 (36)
Ever used hormone therapy	46 (25)	82 (29)	25 (26)	47 (23)	38 (20)	82 (28)
Adiposity measures						
Waist circumference - cm						
≤80	64 (34)	113 (39)	29 (30)	82 (41)	64 (33)	114 (40)
(80-88]	42 (22)	71 (25)	17 (17)	46 (23)	36 (19)	72 (25)
>88	81 (43)	103 (36)	52 (53)	74 (37)	93 (48)	102 (35)
Waist circumference - cm , median (IQR)	86.5 (77.5-96.2)	84.0 (76.0-93.0)	89.8 (78.8-101.3)	83.3 (76.0-93.0)	87.0 (78.0-98.5)	83.8 (76.0-92.5)
Body mass index - kg/m2						
[18.5-25)	58 (31)	103 (36)	33 (34)	75 (37)	58 (30)	103 (36)
[25-30)	69 (37)	109 (38)	24 (24)	75 (37)	66 (34)	110 (38)
≥30	60 (32)	75 (26)	41 (42)	52 (26)	69 (36)	75 (26)
Body mass index - kg/m2, median (IQR)	26.9 (24.0-31.1)	26.7 (23.7-30.3)	28.2 (24.1-34.8)	26.6 (23.5-30.2)	27.5 (24.6-31.3)	26.7 (23.8-30.3)
Waist-hip ratio						
≤0.77	41 (22)	85 (30)	27 (28)	62 (31)	47 (24)	83 (29)
(0.77-0.83]	65 (35)	107 (37)	27 (28)	72 (36)	57 (30)	111 (39)
>0.83	81 (43)	95 (33)	44 (45)	68 (34)	89 (46)	94 (33)
Waist-hip ratio, median (IQR)	0.8 (0.8-0.9)	0.8 (0.8-0.9)	0.8 (0.8-0.9)	0.8 (0.8-0.9)	0.8 (0.8-0.9)	0.8 (0.8-0.9)

continued

Table 8-3 continued Baseline characteristics of women included in the analytic dataset for mediation analysis

	ER-positive Breast Cancer		Endometrial Cancer		Colorectal Cancer	
Biomarkers, median (IQR)						
Adiponectin - µg/mL	27.90 (19.92-37.85)	28.17 (20.20-39.64)	26.62 (16.68-39.89)	28.57 (20.42-40.78)	27.46 (18.63-39.51)	28.24 (20.31-39.58)
Leptin - ng/mL	16.87 (9.22-29.76)	14.85 (7.60-26.34)	20.28 (8.43-38.65)	14.72 (7.58-24.64)	16.92 (8.63-30.28)	14.83 (7.61-26.06)
Resistin - ng/mL	11.3 (9.3-14.7)	12.2 (9.4-15.5)	12.4 (10.5-16.6)	12.2 (9.3-15.0)	13.0 (10.2-16.7)	12.1 (9.4-15.4)
HGF- pg/mL	660.2 (463.8-1000.1)	636.0 (426.1-894.3)	774.7 (466.1-1083.4)	642.4 (420.5-872.8)	674.4 (486.9-990.7)	639.8 (425.7-899.7)
PAI-1 - ng/mL	17.46 (12.42-26.35)	15.74 (10.86-23.22)	17.33 (12.08-27.12)	15.67 (10.37-23.43)	18.98 (12.98-27.51)	15.78 (10.84-23.09)
TNF-α- pg/mL	2.6 (1.9-3.7)	2.8 (1.9-3.7)	2.8 (1.8-4.2)	2.7 (1.8-3.6)	3.0 (1.9-4.0)	2.8 (1.9-3.8)
IL-6 - pg/mL	1.5 (0.9-2.3)	1.5 (1.0-2.4)	1.4 (1.0-2.4)	1.6 (1.0-2.4)	1.7 (1.0-2.7)	1.5 (1.0-2.4)
C-reactive protein - ug/mL	1.8 (0.9-3.6)	1.4 (0.6-3.3)	2.2 (1.1-4.7)	1.4 (0.6-3.3)	2.0 (0.8-4.0)	1.4 (0.6-3.3)
Insulin - uIU/mL	6.5 (4.4-10.6)	5.5 (3.7-8.7)	6.5 (4.4-10.3)	5.2 (3.6-8.1)	7.1 (4.2-11.0)	5.4 (3.7-8.6)
Estradiol - pg/mL	13.0 (8.3-17.0)	11.0 (6.7-16.0)	13.4 (9.1-21.0)	11.0 (6.5-15.0)	11.0 (7.6-16.0)	11.0 (6.7-16.0)

Abbreviations: ER Estrogen Receptor; IQR interquartile range; HGF Hepatocyte growth factor; PAI-1 Plasminogen activator inhibitor-1; TNF-α Tumour necrosis factor-α; IL-6 Interleukin-6; CRP C-reactive protein; µg/mL micrograms per millilitre; ng/mL nanograms per millilitre; pg/mL picograms per millilitre; uIU/mL micro-IU per millilitre

Adiposity-biomarker associations: For waist circumference >88 vs ≤80 cm, adiponectin showed an inverse association (adjusted GMR 0.68; 95%CI, 0.56–0.81), while leptin, resistin, PAI-1, IL-6, and CRP showed positive associations. Of these, the strongest association was observed for leptin (GMR 3.42; 95%CI, 2.75–4.25). Positive associations were also observed for fasting insulin (GMR 2.02; 95%CI, 1.73–2.35) and estradiol (GMR 1.31; 95%CI, 1.09–1.58). The direction and strength of associations were similar for BMI and waist-hip ratio (Table 8-4).

Table 8-4 Adiposity-biomarker association; analysis limited to sub-cohort members who met the eligibility criteria and excluding women categorised as overweight

	Ratio of geometric means (95% confidence interval)								
	Waist circumference >88 vs. ≤80			Body Mass Index ≥30 vs. <25 kg/m2			Waist-hip ratio >0.83 vs. ≤0.77		
Biomarker	N=218			N=181			N=180		
Adiponectin	0.68	(0.56 to	0.81)	0.68	(0.55 to	0.85)	0.67	(0.55 to	0.81)
Leptin	3.42	(2.75 to	4.25)	4.35	(3.44 to	5.50)	1.88	(1.40 to	2.51)
Resistin	1.10	(1.00 to	1.21)	1.09	(0.98 to	1.21)	0.99	(0.89 to	1.11)
HGF	1.18	(0.96 to	1.46)	1.26	(0.97 to	1.63)	1.06	(0.79 to	1.41)
PAI-1	1.58	(1.36 to	1.84)	1.62	(1.37 to	1.92)	1.41	(1.19 to	1.68)
TNF-alpha	1.15	(0.87 to	1.53)	1.18	(0.86 to	1.62)	1.09	(0.78 to	1.52)
IL-6	1.73	(1.43 to	2.09)	1.72	(1.39 to	2.15)	1.54	(1.24 to	1.89)
CRP	2.57	(1.83 to	3.61)	3.19	(2.16 to	4.72)	2.15	(1.43 to	3.24)
Fasting Insulin	2.02	(1.73 to	2.35)	2.12	(1.79 to	2.52)	1.75	(1.46 to	2.11)
Estradiol	1.31	(1.09 to	1.58)	1.39	(1.13 to	1.71)	1.10	(0.88 to	1.37)

Abbreviations: HGF Hepatocyte growth factor; PAI-1 Plasminogen activator inhibitor-1; TNF-α Tumour necrosis factor-α; IL-6 Interleukin-6; CRP C-reactive protein. All analyses were complete case. Models were adjusted for age at baseline, educational attainment, alcohol intake, ever smoked, physical activity, NSAID use, age at menarche, parity, oral contraceptive use, hormone therapy use.

Adiposity-cancer and biomarker-cancer associations: For ER-positive breast cancer, an increased HR was observed for women with waist circumference >88 vs ≤80 cm (adjusted HR 1.61; 95%CI, 1.00–2.59), BMI ≥30 *vs*. <25 kg/m^2 (HR 2.24; 95%CI, 1.24–4.06), and waist-hip ratio >0.83 vs ≤0.77 (HR 1.83; 95%CI, 1.01–3.29). There were positive associations with leptin, CRP, fasting insulin, and estradiol, with the strongest association observed for fasting insulin (HR adjusted for confounders, waist circumference and CRP, per doubling fasting insulin concentration 1.61; 95%CI, 1.23–2.11) (Table 8-5).

For endometrial cancer, a positive association was observed with waist circumference >88 vs ≤80 cm (adjusted HR 2.25; 95%CI, 1.22–4.14) and BMI ≥30 vs <25 kg/m^2 (HR 2.49; 95%CI, 1.24–5.00). An increased HR was observed for leptin, HGF, CRP, fasting insulin, and estradiol, with the strongest association for estradiol (HR adjusted for confounders, waist circumference, CRP, and insulin, per doubling estradiol concentration 1.76; 95%CI, 1.19–2.59) (Table 8-5).

For colorectal cancer, the adjusted HR was 1.51 (95%CI, 0.95–2.39) for waist circumference >88 vs ≤80 cm and 1.89 (95%CI, 1.08–3.30) for BMI ≥30 vs <25 kg/m^2.

Of the biomarkers for which there was evidence for a positive association (i.e. resistin, PAI-1, and fasting insulin), resistin showed the strongest association (HR adjusted for confounders and waist circumference, 1.57; 95%CI, 1.01–2.43) (Table 8-5).

For all cancer sites, the HRs and 95%CIs estimated from the Cox regression models were similar to the RRs and 95%CIs from the logistic regression models. There was no strong evidence for departure from linearity for any of the biomarkers (Table 8-5). We did not observe an association between any of the adiposity measures and TNF-α (Table 8-4), nor between TNF-α and any of the cancers (Table 8-5); TNF-α was thus not included in the mediation analyses.

Table 8-5 Exposure-outcome and mediator-outcome associations

	ER-positive Breast Cancer			Endometrial Cancer			Colorectal Cancer		
	HR (95% CI)	**RR (95% CI)**		**HR (95% CI)**	**RR (95% CI)**		**HR (95% CI)**	**RR (95% CI)**	
Waist Circumference	145 cases, 232 controls	145 cases, 216 controls		81 cases, 163 controls	81 cases, 156 controls		157 cases, 233 controls	157 cases, 216 controls	
>88 vs ≤80	1.61 (1.00 to 2.59)	1.63 (1.03 to 2.58)		2.25 (1.22 to 4.14)	2.26 (1.24 to 4.11)		1.51 (0.95 to 2.39)	1.57 (1.00 to 2.46)	
Body Mass Index	118 cases, 191 controls	118 cases, 178 controls		74 cases, 133 controls	74 cases, 127 controls		127 cases, 193 controls	127 cases, 178 controls	
≥30 vs ≥18.5-<25	2.24 (1.24 to 4.06)	2.20 (1.28 to 3.76)		2.49 (1.24 to 5.00)	2.38 (1.24 to 4.60)		1.89 (1.08 to 3.30)	1.71 (1.03 to 2.85)	
Waist-Hip ratio	122 cases, 194 controls	122 cases, 180 controls		71 cases, 137 controls	71 cases, 130 controls		136 cases, 192 controls	136 cases, 177 controls	
>0.83 vs ≤0.77	1.83 (1.01 to 3.29)	1.88 (1.09 to 3.22)		1.41 (0.70 to 2.81)	1.49 (0.78 to 2.83)		1.35 (0.79 to 2.28)	1.54 (0.93 to 2.56)	
Biomarker	**Per doubling concentration**		P value from likelihood ratio test for evidence against linearity*	**Per doubling concentration**		P value from likelihood ratio test for evidence against linearity*	**Per doubling concentration**		P value from likelihood ratio test for evidence against linearity*
	HR (95% CI)	RR (95% CI)		HR (95% CI)	RR (95% CI)		HR (95% CI)	RR (95% CI)	
	187 cases, 307 controls	187 cases, 287 controls		98 cases, 213 controls	98 cases, 202 controls		193 cases, 310 controls	193 cases, 288 controls	
Adiponectin									
Model 1	0.98 (0.78 to 1.24)	1.02 (0.82 to 1.27)	0.59	0.82 (0.61 to 1.10)	0.86 (0.66 to 1.13)	0.70	0.92 (0.73 to 1.17)	0.96 (0.78 to 1.20)	0.11
Leptin									
Model 1	1.30 (1.05 to 1.60)	1.27 (1.06 to 1.52)	0.38	1.36 (1.03 to 1.79)	1.28 (1.00 to 1.64)	0.05	1.17 (0.97 to 1.42)	1.12 (0.94 to 1.34)	0.70
Resistin									
Model 1	0.96 (0.65 to 1.44)	0.96 (0.64 to 1.42)	0.51	1.31 (0.75 to 2.28)	1.29 (0.77 to 2.16)	0.87	1.57 (1.01 to 2.43)	1.42 (0.97 to 2.07)	0.35
HGF									
Model 1	1.13 (0.92 to 1.37)	1.08 (0.91 to 1.29)	0.42	1.26 (1.02 to 1.57)	1.29 (0.99 to 1.68)	0.07	1.02 (0.85 to 1.23)	1.02 (0.87 to 1.19)	0.39
PAI-1									
Model 1	1.28 (0.95 to 1.72)	1.26 (0.96 to 1.65)	0.36	1.26 (0.87 to 1.81)	1.19 (0.85 to 1.67)	0.63	1.49 (1.10 to 2.02)	1.35 (1.03 to 1.76)	0.22

continued

Table 8-5 continued Exposure-outcome and mediator-outcome associations

Biomarker	Per doubling concentration HR (95% CI) 187 cases, 307 controls	Per doubling concentration RR (95% CI) 187 cases, 287 controls	P value from likelihood ratio test for evidence against linearity*	Per doubling concentration HR (95% CI) 98 cases, 213 controls	Per doubling concentration RR (95% CI) 98 cases, 202 controls	P value from likelihood ratio test for evidence against linearity*	Per doubling concentration HR (95% CI) 193 cases, 310 controls	Per doubling concentration RR (95% CI) 193 cases, 288 controls	P value from likelihood ratio test for evidence against linearity*
TNF-α									
Model 1	1.07 (0.93 to 1.24)	1.07 (0.92 to 1.24)	0.75	0.99 (0.82 to 1.20)	0.99 (0.84 to 1.16)	0.68	1.00 (0.85 to 1.18)	0.98 (0.87 to 1.12)	0.16
Il-6									
Model 1	0.90 (0.72 to 1.12)	0.87 (0.71 to 1.08)	0.75	0.88 (0.66 to 1.18)	0.88 (0.67 to 1.17)	0.91	1.02 (0.81 to 1.29)	0.95 (0.78 to 1.17)	0.07
CRP									
Model 1	1.12 (1.00 to 1.26)	1.16 (1.02 to 1.32)	0.06	1.20 (1.02 to 1.40)	1.19 (1.01 to 1.41)	0.63	1.06 (0.95 to 1.20)	1.04 (0.92 to 1.18)	0.72
Fasting insulin									
Model 1	1.65 (1.26 to 2.16)	1.67 (1.27 to 2.19)		1.60 (1.07 to 2.38)	1.42 (1.01 to 2.00)		1.42 (1.09 to 1.86)	1.35 (1.05 to 1.74)	
Model 2 (add adj for CRP)	1.61 (1.23 to 2.11)	1.61 (1.22 to 2.12)	0.59	1.50 (1.01 to 2.24)	1.34 (0.95 to 1.90)	0.55	1.41 (1.08 to 1.84)	1.34 (1.03 to 1.74)	0.38
Estradiol									
Model 1	1.26 (1.03 to 1.53)	1.26 (1.02 to 1.56)		1.76 (1.27 to 2.44)	1.73 (1.29 to 2.33)		1.01 (0.82 to 1.25)	1.03 (0.83 to 1.28)	
Model 2 (add adj for CRP, fasting insulin)	1.24 (1.00 to 1.54)	1.22 (0.97 to 1.53)	0.31	1.76 (1.19 to 2.59)	1.68 (1.24 to 2.27)	0.59	0.99 (0.79 to 1.24)	1.02 (0.81 to 1.27)	0.32

Abbreviations: HR Hazard Ratio; CI Confidence Interval; RR Risk Ratio; HGF Hepatocyte growth factor; PAI-1 Plasminogen activator inhibitor-1; TNF-α Tumour necrosis factor-α; IL-6 Interleukin-6; CRP C-reactive protein. All analyses were complete case. Hazard ratios were estimated from Cox models that allowed for the case-cohort design using the Prentice approach and for lost to follow-up (i.e. women were censored at the time of death, lost to follow-up, or end of follow-up, whichever occurred first). Risk ratios were estimated from logistic regression models that excluded women who died or were lost to follow up before the end of follow-up (29th February 2004) of this study. Models were adjusted for age at baseline, educational attainment, alcohol intake, ever smoked, physical activity, NSAID use, age at menarche, parity, oral contraceptive use, hormone therapy use. * Reported P values are from the Cox regression models. The number of participants included in the Cox regression model and logistic regression model are different because for the latter, women who died or were lost to follow-up before the end of the follow-up period were excluded.

Mediation analysis:

For ER-positive breast cancer, the RR^{IIE} comparing women with waist circumference >88 vs ≤80 cm was 2.09 (95%CI, 1.15–3.46) through all biomarkers, 1.10 (95%CI, 0.61–1.96) through inflammatory status, 1.67 (95%CI, 1.08–2.65) through fasting insulin, and 1.14 (95%CI, 1.02–1.39) through estradiol. The RR^{IDE} not explained by the included biomarkers was 0.80 (95%CI, 0.39–1.71). The patterns of mediation were comparable for BMI and WHR (Table 8-6).

For endometrial cancer, the RR^{IIE} for women with waist circumference >88 vs ≤80 cm was 2.01 (95%CI, 0.86–3.97) through all biomarkers, 1.40 (95%CI, 0.58-3.06) through inflammatory status, 1.13 (95%CI, 0.61–1.83) through fasting insulin, and 1.27 (95%CI, 1.04–1.69) through estradiol. The RR^{IDE} not explained by the biomarkers was 1.11 (95%CI, 0.43–3.75). Comparable results were observed when BMI was used the adiposity measure (Table 8-6).

For colorectal cancer, the RR^{IIE} for women with waist circumference >88 vs ≤80 cm was 1.61 (95%CI, 0.95–2.52) through all biomarkers, 1.37 (95%CI, 0.82–2.38) through inflammatory status, 1.16 (95%CI, 0.77–1.69) through fasting insulin, and 1.01 (95%CI, 0.89–1.14) through estradiol. The RR^{IDE} not explained by the biomarkers was 1.00 (95%CI, 0.51–1.94). The patterns of mediation were comparable for BMI (Table 8-6).

Table 8-6 Interventional indirect and direct effects

		Waist circumference >88 vs ≤80 cm		BMI ≥30 vs ≥18.5-<25 kg/m2		Waist-hip ratio >0.83 vs ≤0.77	
	Effects	Risk Ratio (95% CI)	Risk Difference per 10,000 (95% CI)	Risk Ratio (95% CI)	Risk Difference per 10,000 (95% CI)	Risk Ratio (95% CI)	Risk Difference per 10,000 (95% CI)
ER-positive Breast Cancer	Total causal effect*	1.68 (0.96 to 2.79)	29.99 (1.35 to 62.19)	2.35 (1.22 to 4.34)	49.88 (15.73 to 101.92)	1.82 (0.93 to 3.35)	32.66 (2.68 to 69.85)
	Interventional indirect effect through all biomarkers	2.09 (1.15 to 3.46)	38.66 (16.45 to 67.28)	3.82 (1.71 to 7.69)	64.15 (38.51 to 114.24)	1.55 (0.93 to 2.70)	25.62 (1.15 to 55.57)
	Interventional indirect effect through inflammatory status	1.10 (0.61 to 1.96)	3.34 (-19.73 to 23.40)	1.70 (0.70 to 3.75)	15.52 (-4.01 to 46.05)	0.89 (0.53 to 1.49)	-5.21 (-41.76 to 14.24)
	Interventional indirect effect through fasting insulin	1.67 (1.08 to 2.65)	25.65 (9.81 to 50.37)	1.87 (1.12 to 3.26)	32.95 (11.29 to 74.30)	1.66 (1.16 to 2.84)	28.58 (10.29 to 70.70)
	Interventional indirect effect through estradiol	1.14 (1.02 to 1.39)	8.78 (1.80 to 24.06)	1.20 (1.02 to 1.60)	14.18 (2.85 to 41.83)	1.04 (0.97 to 1.24)	2.71 (-2.32 to 17.97)
	Interventional direct effect	0.80 (0.39 to 1.71)	-8.68 (-32.94 to 21.76)	0.62 (0.24 to 1.61)	-14.28 (-36.37 to 10.33)	1.18 (0.52 to 2.50)	7.03 (-21.27 to 44.27)
	Interdependence	1.00 (1.00 to 1.03)	0.89 (-1.68 to 5.46)	1.00 (0.99 to 1.01)	1.50 (-5.05 to 11.79)	1.00 (1.00 to 1.07)	-0.45 (-10.56 to 2.82)

continued

Table 8-6 continued Interventional indirect and direct effects

		Waist circumference >88 vs ≤80 cm		BMI ≥30 vs ≥18.5-<25 kg/m2		Waist-hip ratio >0.83 vs ≤0.77	
	Effects	Risk Ratio (95% CI)	Risk Difference per 10,000 (95% CI)	Risk Ratio (95% CI)	Risk Difference per 10,000 (95% CI)	Risk Ratio (95% CI)	Risk Difference per 10,000 (95% CI)
Endometrial Cancer	Total causal effect*	2.24 (1.03 to 5.03)	36.67 (9.17 to 82.91)	2.38 (0.97 to 5.44)	48.25 (10.56 to 122.02)		
	Interventional indirect effect through all biomarkers	2.01 (0.86 to 3.97)	33.36 (6.61 to 67.52)	3.16 (1.04 to 8.26)	56.99 (23.81 to 116.94)		
	Interventional indirect effect through inflammatory status	1.40 (0.58 to 3.06)	13.14 (-15.43 to 54.78)	1.99 (0.61 to 5.39)	25.09 (-5.33 to 91.28)		
	Interventional indirect effect through fasting insulin	1.13 (0.61 to 1.83)	5.87 (-28.01 to 25.47)	1.22 (0.60 to 2.32)	10.84 (-36.35 to 42.87)		
	Interventional indirect effect through estradiol	1.27 (1.04 to 1.69)	13.97 (3.89 to 37.51)	1.30 (0.99 to 2.05)	18.53 (0.65 to 57.19)		
	Interventional direct effect	1.11 (0.43 to 3.75)	3.30 (-19.47 to 55.14)	0.75 (0.21 to 3.11)	-8.74 (-36.57 to 42.08)		
	Interdependence	1.00 (1.00 to 1.00)	0.38 (-2.60 to 10.73)	1.00 (0.99 to 1.00)	2.53 (-5.37 to 22.21)		

continued

Table 8-6 continued Interventional indirect and direct effects

		Waist circumference >88 vs ≤80 cm		BMI ≥30 vs ≥18.5-<25 kg/m2		Waist-hip ratio >0.83 vs ≤0.77	
	Effects	Risk Ratio (95% CI)	Risk Difference per 10,000 (95% CI)	Risk Ratio (95% CI)	Risk Difference per 10,000 (95% CI)	Risk Ratio (95% CI)	Risk Difference per 10,000 (95% CI)
Colorectal Cancer	Total causal effect*	1.61 (0.94 to 2.58)	29.23 (1.46 to 64.34)	1.76 (0.95 to 3.06)	35.57 (2.29 to 77.44)		
	Interventional indirect effect through all biomarkers	1.61 (0.95 to 2.52)	29.06 (2.12 to 56.70)	2.02 (1.06 to 4.00)	41.64 (12.76 to 80.24)		
	Interventional indirect effect through inflammatory status	1.37 (0.82 to 2.38)	17.32 (-7.27 to 55.71)	1.36 (0.69 to 2.87)	14.31 (-14.56 to 57.57)		
	Interventional indirect effect through fasting insulin	1.16 (0.77 to 1.69)	10.51 (-21.40 to 33.17)	1.50 (0.88 to 2.39)	27.21 (-1.46 to 62.38)		
	Interventional indirect effect through estradiol	1.01 (0.89 to 1.14)	0.91 (-9.67 to 9.18)	0.99 (0.83 to 1.17)	-0.60 (-17.25 to 12.16)		
	Interventional direct effect	1.00 (0.51 to 1.94)	0.18 (-26.36 to 35.82)	0.87 (0.39 to 2.03)	-6.07 (-33.73 to 32.90)		
	Interdependence	1.00 (0.99 to 1.00)	0.32 (-0.37 to 2.48)	1.00 (0.98 to 1.00)	0.73 (-0.47 to 6.99)		

*Total effects presented here were estimated as the product of interventional indirect and direct effects. All analyses were complete case. Models were adjusted for age at baseline, educational attainment, alcohol intake, ever smoked, physical activity, NSAID use, age at menarche, parity, oral contraceptive use, hormone therapy use.

Additional analyses:

In the additional analyses that estimated the RRs^{IIE} through each of the adipokines and through IL-6 and CRP jointly, the point estimates were suggestive of the strongest IIE for ER-positive breast cancer through leptin and PAI (RR^{IIE} for waist circumference >88 vs ≤80 cm 1.11; 95%CI, 0.69–1.76 and 1.20; 95%CI, 0.91–1.57 respectively), for endometrial cancer through leptin (RR^{IIE} for waist circumference >88 vs ≤80 cm 1.38; 95%CI, 0.73–2.79), and for colorectal cancer for leptin and PAI (RR^{IIE} for waist circumference >88 vs ≤80 cm 1.22; 95%CI, 0.78–1.92 and 1.22; 95%CI, 0.97–1.61 respectively) (Supplementary Table D-3).

Discussion

In this study of postmenopausal women, we quantified the mediating effects of inflammatory status, fasting insulin, and estradiol in the associations between adiposity and ER-positive breast, endometrial and colorectal cancers. For each cancer site, the analysis suggested that the assessed biomarkers explained the observed associations to a large extent. However, there was high uncertainty around the estimates of the total, indirect, and direct effects, making definitive judgments about the magnitude of the mediating effects difficult. The results also indicated that the relative importance of the assessed biomarkers might be different for the three cancer sites. For ER-positive breast cancer, the largest indirect effect was observed through fasting insulin, followed by estradiol. For endometrial cancer, the point estimate through inflammatory status was largest, followed by estradiol, but the 95%CI around the estimated indirect effect through inflammatory status was wide. For colorectal cancer, the 95%CIs around all the estimated effects were wide but depending on the measure of adiposity the largest indirect effect point estimate was observed for inflammatory status (waist circumference) or insulin (BMI), while the effect through estradiol was essentially null.

To our knowledge, this is one of the first studies that has attempted to quantify the mediating role of inflammatory status, hyperinsulinemia, and estradiol in explaining the effect of adiposity on the risk of ER-positive breast, endometrial, and colorectal cancer in postmenopausal women using data from the same cohort. The causal mediation analysis approach used allowed us to appropriately account for the potential dependencies between the assessed biomarkers. Also, the interventional indirect and direct effects can be used without needing a prespecified causal ordering of the mediators, which compared with other causal mediation analysis approaches that require such assumption (for example the sequential causal mediation analysis approach described in reference (129)) is a considerable advantage.

The estimated mediating roles for the three pathways would have depended on how well the measured biomarkers captured the biological characteristics of each pathway, the quality of the measurements, and the temporal stability of the biomarkers. We had measures for a range of adipocytokines and cytokines to capture the inflammatory status. However, for the sex-steroid hormone pathway, it is unlikely that estradiol on its own would have reflected all aspects of the pathway. In general, the biomarker measurements were of good quality (for adipocytokines and cytokines the inter-assay coefficients of variations (CV) were ≤13% (59, 68, 76), estradiol and insulin tests were performed in duplicates and tests with respectively >20% and >10% CV were repeated (83, 90, 119)). For all biomarkers, we had measurements for a single time point. Previously published data on the intraclass correlation coefficient (ICC) calculated using data collected about three years apart for the seven adipocytokines and cytokines used here (195, 196) and estradiol (197) suggest that the single time point measurements might have reflected the average levels over time (all ICCs were >40%).

We excluded cases diagnosed within the first year after baseline to reduce the possibility of cancer having influenced measures of adiposity or biomarkers. However, we cannot rule out

reserve causation for the adiposity-biomarker associations, because they were measured cross-sectionally at baseline.

An important limitation was the small number of cases and controls for each cancer site. Due to small sample size, we were also not able to include any adiposity-biomarker or biomarker-biomarker interaction terms in our mediation analysis models, which might have led to a misspecification bias in the estimation of the indirect and direct effects (155).

Using data from two case-cohort studies within the WHI-OS, one of which was the same study used here, a causal mediation analysis was previously performed to investigate the mediating roles of estradiol and fasting insulin in the effect of BMI on postmenopausal breast cancer in women not using hormone therapy (120). Of the total BMI-breast cancer association (additional cases per 100,000 per 5 kg/m^2 increment in BMI 50.0; 95%CI, 23.2–76.6), 24% was mediated through estradiol (indirect effect additional cases per 100,000 11.9, 95%CI, 1.7–22.6), with a larger mediating effect (49%) observed for ER-positive breast cancers. In women without diabetes, of the total effect (additional cases per 100,000 52.0; 95%CI, 12.1–91.3), 66% was mediated through fasting insulin (indirect effect additional cases per 100,000 34.2; 94%CI, 9.4–59.0) (120). Results from our study cannot be directly compared with these because the effects were estimated on different scales. However, both suggest that fasting insulin and estradiol explained some of the adiposity-postmenopausal breast cancer association, with fasting insulin having a larger role. Contrary to the previous study that assumed independence between estradiol and fasting insulin, our analysis allowed for potential dependence between the three assessed pathways. In another causal mediation analysis of a case-cohort study within the Melbourne Collaborative Cohort Study (MCCS), of the adiposity-ER positive postmenopausal breast cancer association (RR for BMI>30 vs ≤25kg/m^2 1.75; 95%CI, 1.05–2.91), 72% was explained by free estradiol (RR^{IIE} 1.56; 95% CI, 1.11–2.19) and

28% by fasting insulin (RR^{IIE} 1.12; 95%CI, 0.68–1.84) (198). The weaker mediating effect observed for estradiol in the present study might be because we did not have a measure for free estradiol, which, compared with total estradiol, is likely a more important mediator of the adiposity-breast cancer association (16). In line with the weaker mediating effect observed for insulin in the MCCS, the reported insulin-postmenopausal breast cancer association was weaker in that study (HR per doubling concentration 1.00; 95%CI, 0.65–1.53), compared with what we observed here and what was reported in a recent Mendelian randomisation (MR) study (OR per genetically determined one standard deviation in insulin level 2.07; 95%CI, 1.32–3.23) (43). It is difficult to speculate what might explain this difference, but different assays used in the MCCS and WHI-OS, different distribution of confounding factors in the studies, and chance (the 95%CIs from MCCS, WHI-OS and the MR study did overlap) might have all been contributing factors.

In the present study, estradiol mediated the effect of adiposity on endometrial cancer risk only in part. Similarly, in two other studies, adjusting for estradiol partially attenuated the BMI-endometrial cancer association (odds ratio (OR) for BMI ≥30 vs <25kg/m^2 from 3.97 (95%CI, 2.54-6.21) to 2.25 (95%CI, 1.33-3.81) (114), and from 2.67 (95%CI, 1.63-4.37) to 2.09 (95%CI, 1.22-3.57) (113))). Increased estrogen levels in the obese state is hypothesized to play the key role in explaining the adiposity-endometrial cancer risk association (112, 190), but these observations suggest that other important pathways are also likely involved. In our study, the point estimates indicated a larger mediating effect through the inflammatory status compared with estradiol, but the wide 95%CI around the indirect effect did not allow drawing conclusions about presence and magnitude of the mediating effect through this pathway.

Three studies in women (59, 83, 117), two of which employed data from the same WHI-OS case-cohort study used here, investigated the potential mediating role of several biomarkers of

inflammatory status (adiponectin, soluble leptin receptor, leptin, and CRP), C-peptide (a biomarker for hyperinsulinemia), and fasting insulin in the waist circumference-colorectal cancer association by comparing models adjusted and unadjusted for the biomarker. These studies showed that inflammatory status biomarkers, except CRP, weakly attenuated the association (59, 83, 117). Also, C-peptide (117) and fasting insulin (59, 83) explained some (~30%) of the association. When assessed together, leptin and fasting insulin explained 50% of the association (59). In the present study, the indirect effect through all the biomarkers had a magnitude as large as the total effect, suggesting that together, the assessed biomarkers played an important mediating role in the adiposity-colorectal cancer association. However, for each indirect effect, the 95%CIs were wide. An interesting finding here was the essentially null indirect effect observed for estradiol. If replicated in future studies with larger sample sizes and with measures for free estradiol, this could suggest that the hypothesis about the potential protective effect of endogenous estrogens in colorectal cancer development, and their role in explaining the stronger BMI-colorectal cancer association observed in men and the higher incidence of colorectal cancer in men (199) might not hold.

In summary, this study offers some clues about the potential mediating roles of inflammatory status, hyperinsulinemia, and estradiol in the adiposity-ER-positive breast, endometrial, and colorectal cancer associations and their varying importance between the three cancers. Studies with larger sample sizes are needed to further these investigations and produce results with more certainty. Ideally, future studies will take advantage of novel causal mediation analysis approaches, similar to the method used in the present study, to appropriately account for dependence between biomarkers.

Financial Support: This work was supported by Australian National Health and Medical Research Council grants (1150591 to SGD and 1104975 to JAS); and Melbourne Research Scholarship (to SGD). MMB is the recipient of an Australian Research Council Discovery Early Career Award (project number DE190101326) funded by the Australian Government.

The work reported in this paper was undertaken by SGD while hosted by the International Agency for Research on Cancer (IARC) and funded partially by a PhD Scholarship by IARC. The WHI program is funded by the NHLBI through contracts N01WH22110, 24152, 32100-2, 32105-6, 32108-9, 32111-13, 32115, 32118-32119, 32122, 42107-26, 42129-32, and 44221. The complete list of WHI centres and investigators can be found online at http://www.whiscience.org/publications/WHI_investigators_shortlist.pdf.

Chapter 9: Discussion

This thesis aimed to quantify the mediating roles of disturbed adipocytokine production and chronic low-grade inflammation, insulin resistance and hyperinsulinemia, and alteration in levels of sex-steroid hormones in the effect of adiposity on the risk of colorectal, estrogen receptor (ER)-positive postmenopausal breast, and endometrial cancers. Data from four prospective studies were analysed, and novel causal mediation analysis approaches were employed that allowed for potential dependencies between the assessed biomarkers. Chapters 5 to 8 provided a detailed description, including findings, of these investigations. This chapter summarises the main findings, provides an overview of the strengths and limitations of the thesis and discusses the findings, their implications, and suggestions for future research.

9.1 Summary of the main findings

Adiposity and colorectal cancer risk in men

In Chapter 5, data from the UK Biobank cohort were used to explain the effect of adiposity on the risk of colorectal cancer in men. All relevant biomarkers available in the study were used, which included C-reactive protein (CRP, a non-specific biomarker of systemic inflammation), glycated haemoglobin (HbA1c, a correlate of insulin resistance (82)), and sex hormone-binding globulin (SHBG) and testosterone (treated jointly to capture, in part, altered sex-steroid hormone levels). Increased levels of CRP, as well as lower levels of SHBG and testosterone, observed in men with obesity compared with normal weight, explained a small part of the effect of adiposity on colorectal cancer risk. There was no evidence for mediation through HbA1c. Due to the sequential nature of the analysis, the estimated effects through HbA1c and SHBG and testosterone excluded the potential influences of the preceding pathways on these biomarker levels (i.e. potential influence of CRP on HbA1c and CRP and HbA1c on SHBG

and testosterone). A large part of the adiposity-colorectal cancer risk association remained unexplained by the biomarkers included in this analysis (Table 9-1).

Adiposity and colorectal cancer risk in postmenopausal women

The analysis of the UK Biobank data, summarised above, was also performed for postmenopausal women (Chapter 5). There was evidence for a small mediating effect through increased CRP levels, but no evidence for mediation through HbA1c (beyond the potential influence of CRP) or SHBG and testosterone (combined, and beyond the potential influences of CRP and HbA1c). As for men, a large part of the effect remained unexplained (captured by the estimated direct effect) was larger than the estimated indirect effects (Table 9-1).

Data from a case-cohort study within the Women's Health Initiative Observational Study (WHI-OS) were used in Chapter 8 to quantify the mediating roles of a panel of biomarkers reflecting the inflammatory status (i.e. adiponectin, leptin, resistin, plasminogen activator inhibitor-1 (PAI-1), hepatocyte growth factor (HGF), tumour necrosis factor-α (TNF-α), interleukin-6 (IL-6) and CRP), fasting insulin, and estradiol in the effect of adiposity on colorectal cancer risk (as well as on ER-positive breast and endometrial cancer risk) in postmenopausal women. For this study, the interventional mediation analysis approach was used. There was high uncertainty around the estimated mediating effects, precluding the possibility of making any definitive judgement about the presence of mediation through the assessed biomarkers. However, the point estimates were suggestive of weak to moderate mediating effects for inflammatory status and fasting insulin, while the indirect effect through estradiol appeared essentially null (Table 9-1).

Table 9-1 Estimated indirect and direct effects for the association between adiposity measures and colorectal cancer risk for men and postmenopausal women

Study	Estimated effects/ scale	Adiposity measure	No. affected with cancer / No. unaffected	Total effect*** Effect size (95%CI)	Indirect effects Effect size (95%CI)			Direct effect Effect size (95%CI)
					Inflammatory status pathway	Insulin signalling pathway	Sex-steroid hormones pathway	
UK Biobank men*	NIEs & NDE/ Risk ratio	Waist circumference >102 vs. 94 cm	878/127,830	1.37 (1.19–1.58)	1.06 (1.01–1.11)	0.99 (0.97–1.01)	1.03 (0.99–1.08)	1.26 (1.09–1.47)
		Body mass index ≥30 vs. <25 kg/m^2	614/88,142	1.36 (1.15–1.60)	1.06 (0.99–1.13)	0.99 (0.96–1.02)	1.05 (0.99–1.12)	1.23 (1.02–1.48)
		Waist hip ratio >0.98 vs. <0.89	600/88,026	1.31 (1.09–1.58)	1.07 (1.00–1.14)	0.99 (0.96–1.02)	1.06 (1.01–1.11)	1.17 (0.95–1.44)
UK Biobank Postmenopausal women*	NIEs & NDE/ Risk ratio	Waist circumference >88 vs. ≤80 cm	616/100,142	1.27 (1.07–1.50)	1.08 (0.99–1.17)	1.00 (0.98–1.02)	1.00 (0.92–1.09)	1.18 (0.96–1.45)
		Body mass index ≥30 vs. <25 kg/m^2	451/82,506	-****	-	-	-	-
		Waist hip ratio >0.86 vs. ≤0.77	378/65,704	1.33 (1.06–1.66)	1.09 (1.00–1.19)	1.01 (0.97–1.05)	1.00 (0.91–1.11)	1.20 (0.92–1.57)
WHI-OS Postmenopausal women**	IIEs & IDE/ Risk ratio	Waist circumference >88 vs. ≤80 cm	157/ 216	1.61 (0.94–2.58)	1.37 (0.82–2.38)	1.16 (0.77–1.69)	1.01 (0.89–1.14)	1.00 (0.51–1.94)
		Body mass index ≥30 vs. <25 kg/m^2	127/ 178	1.76 (0.95–3.06)	1.36 (0.69–2.87)	1.50 (0.88–2.39)	0.99 (0.83–1.17)	0.87 (0.39–2.03)
		Waist hip ratio >0.83 vs ≤0.77	136/ 177	-****	-	-	-	-

Abbreviations WHI-OS Women's Health Initiative Observational Study; NIE natural indirect effect; NDE natural direct effect; IIE interventional indirect effect; IDE interventional direct effect; CI confidence interval

*In the UK Biobank, assessed biomarkers on each pathway were: inflammatory status: C-reactive protein; insulin signalling: glycated haemoglobin; sex-steroid hormones: sex hormone binding globulin and testosterone treated jointly

**In the WHI-OS study, assessed biomarkers on each pathway were: inflammatory status: adiponectin, leptin, resistin, plasminogen activator inhibitor-1, hepatocyte growth factor, interleukin-6, and C-reactive protein; insulin signalling: fasting insulin; sex-steroid hormones: total estradiol

***All Total effects presented in the table were estimated as the product of indirect and direct effects.

****Mediation analysis was not performed because the 95% confidence interval around the total effect estimated from the regression model for the outcome conditional on exposure and baseline confounders was wide and included both decreased and increased risks.

Adiposity and ER-positive breast cancer risk in postmenopausal women

In Chapter 6, data from a case-cohort study within the Melbourne Collaborative Cohort Study (MCCS) were analysed using the interventional mediation analysis approach to quantify the mediating roles of fasting insulin and calculated free estradiol in the effect of adiposity on ER-positive breast cancer risk in postmenopausal women. No biomarkers reflecting inflammatory status were measured in this study. A moderate to strong mediation was observed through increased free estradiol levels. Although the point estimates were also suggestive of a weak mediation through fasting insulin, there was high uncertainty around the estimated indirect effects, with the 95% confidence intervals including both decreased and increased risks (Table 9-2).

The analysis of the WHI-OS data (Chapter 8), described above, was also performed for ER-positive breast cancer risk in postmenopausal women. A strong mediating effect was observed through fasting insulin and a weak to moderate mediation for estradiol. For inflammatory status, there was high uncertainty around the estimated indirect effects and the point estimates for different measures of adiposity were inconsistent, suggesting a weak (waist circumference), moderate (body mass index (BMI)), or even an inverse (waist-hip ratio) mediating effect (Table 9-2).

Table 9-2 Estimated indirect and direct effects for the association between adiposity measures and estrogen receptor positive breast cancer for postmenopausal women

Study	Estimated effects/ scale	Adiposity measure	No. affected with cancer / No. unaffected	Total effect*** Effect size (95%CI)	Indirect effects Effect size (95%CI)			Direct effect Effect size (95%CI)
					Inflammatory status pathway	Insulin signalling pathway	Sex-steroid hormones pathway	
MCCS Postmenopausal women*	IIEs & IDE/ Risk ratio	Waist circumference >88 vs. ≤80 cm	117/764	1.96 (1.23–3.13)	-	1.23 (0.81–1.88)	1.42 (1.09–1.87)	1.01 (0.81–1.25)
		Body mass index ≥30 vs. <25 kg/m^2	97/604	1.75 (1.05–2.91)	-	1.12 (0.68–1.84)	1.56 (1.11–2.19)	1.03 (0.43–2.48)
WHI-OS Postmenopausal women**	IIEs & IDE/ Risk ratio	Waist circumference >88 vs. ≤80 cm	145/216	1.68 (0.96–2.79)****	1.10 (0.61–1.96)	1.67 (1.08–2.65)	1.14 (1.02–1.39)	0.80 (0.39–1.71)
		Body mass index ≥30 vs. <25 kg/m^2	118/178	2.35 (1.22–4.34)	1.70 (0.70–3.75)	1.87 (1.12–3.26)	1.20 (1.02–1.60)	0.62 (0.24–1.61)
		Waist hip ratio >0.83 vs ≤0.77	122/180	1.82 (0.93–3.35)****	0.89 (0.53–1.49)	1.66 (1.16–2.84)	1.04 (0.97–1.24)	1.18 (0.52–2.50)

Abbreviations MCCS Melbourne Collaborative Cohort Study; WHI-OS Women's Health Initiative Observational Study; IIE interventional indirect effect; IDE interventional direct effect; CI confidence interval

*In the MCCS, assessed biomarkers on each pathway were: insulin signalling: fasting insulin; sex-steroid hormones: calculated free estradiol. Inflammatory status pathway was not assessed. Association with waist-hip ratio was not report in the MCCS manuscript.

**In the WHI-OS study, assessed biomarkers on each pathway were: inflammatory status: adiponectin, leptin, resistin, plasminogen activator inhibitor-1, hepatocyte growth factor, interleukin-6, and C-reactive protein; insulin signalling: fasting insulin; sex-steroid hormones: total estradiol

***All Total effects presented in the table were estimated as the product of indirect and direct effects.

****Although the 95% confidence intervals around the total effect presented here, which was estimated as the product of interventional indirect and direct effects, are wide and include both decreased and increased risks, mediation analysis was still performed because based on the regression model for the outcome conditional on exposure and baseline confounders there was evidence for a total effect.

Adiposity and endometrial cancer risk in postmenopausal women

In Chapter 7, data from a case-control study nested within the European Prospective Investigation into Cancer and Nutrition (EPIC) were used to quantify the mediating roles of a panel of adipocytokines and cytokines to capture the inflammatory status (adiponectin, interleukin-6 (IL-6), interleukin-1-receptor antagonist (IL-1Ra), TNF-receptor (TNF-R)-1 and -2, and CRP), C-peptide (hyperinsulinemia biomarker), and free estradiol and estrone (treated jointly to capture altered sex-steroid hormone levels) on adiposity-endometrial cancer link in postmenopausal women. Sequential mediation analysis was used. There was some evidence for a moderate mediation through inflammatory status. The point estimates were also suggestive of weak mediation through C-peptide (beyond the potential influence of inflammatory status on C-peptide levels) and weak to moderate mediation through free estradiol and estrone (beyond the influences of inflammatory status and C-peptide), but there was high uncertainty around the estimated indirect effects (Table 9-3).

The analysis of the WHI-OS data (Chapter 8), described above, was also performed for endometrial cancer risk in postmenopausal women. A moderate mediating effect through estradiol was observed (Table 9-3). The point estimates were also suggestive of moderate mediation for inflammatory status and weak to moderate mediation though fasting insulin, but again, there was high uncertainty around the estimated indirect effects.

Table 9-3 Estimated indirect and direct effects for the association between adiposity measures and endometrial cancer for postmenopausal women

Study	Estimated effects/ scale	Adiposity measure	No. affected with cancer / No. unaffected	Total effect*** Effect size (95%CI)	Indirect effects Effect size (95%CI)			Direct effect Effect size (95%CI)
					Inflammatory status pathway	Insulin signalling pathway	Sex-steroid hormones pathway	
EPIC Postmenopausal women*	NIEs & NDE/ Odds ratio	Waist circumference >88 vs. ≤80 cm	131/212	2.07 (1.20–3.55)	1.56 (1.01–2.42)	1.03 (0.89–1.19)	1.08 (0.88–1.33)	1.19 (0.59–2.41)
		Body mass index ≥30 vs. <25 kg/m^2	102/185	2.51 (1.26–5.02)	1.53 (0.89–2.62)	1.05 (0.88–1.24)	1.22 (0.89–1.67)	1.29 (0.54–3.09)
WHI-OS Postmenopausal women**	IIEs & IDE/ Risk ratio	Waist circumference >88 vs. ≤80 cm	81/156	2.24 (1.03–5.03)	1.40 (0.58–3.06)	1.13 (0.61–1.83)	1.27 (1.04–1.69)	1.11 (0.43–3.75)
		Body mass index ≥30 vs. <25 kg/m^2	74/127	2.38 (0.97–5.44)****	1.99 (0.61–5.39)	1.22 (0.60–2.32)	1.30 (0.99–2.05)	0.75 (0.21–3.11)
		Waist hip ratio >0.83 vs ≤0.77	71/130	- †	-	-	-	-

Abbreviations EPIC European Prospective Investigation into Cancer and Nutrition; WHI-OS Women's Health Initiative Observational Study; NIE natural indirect effect; NDE natural direct effect; IIE interventional indirect effect; IDE interventional direct effect; CI confidence interval

*In the EPIC, assessed biomarkers on each pathway assessed biomarkers on each pathway were: inflammatory status: interleukin-6, interleukin-1 receptor antagonist, tumour necrosis factor-a, TNF receptor 1, TNFR-2, and C-reactive protein treated jointly; insulin signalling: C-peptide; sex-steroid hormones: calculated free estradiol and estrone treated jointly. Association with waist-hip ratio was not report in the MCCS manuscript.

**In the WHI-OS study, assessed biomarkers on each pathway were: inflammatory status: adiponectin, leptin, resistin, plasminogen activator inhibitor-1, hepatocyte growth factor, interleukin-6, and C-reactive protein; insulin signalling: fasting insulin; sex-steroid hormones: total estradiol

*** All Total effects presented in the table were estimated as the product of indirect and direct effects.

**** Although the 95% confidence intervals around the total effect presented here, which was estimated as the product of interventional indirect and direct effects, are wide and include both decreased and increased risks, mediation analysis was still performed because based on the regression model for the outcome conditional on exposure and baseline confounders there was evidence for a total effect.

† Mediation analysis was not performed because the 95% confidence interval around the total effect was wide and included both decreased and increased risks.

Comparison of the mediating role of pathways between the three cancers

The analysis of the WHI-OS case-cohort study of colorectal, ER-positive breast, and endometrial cancer in postmenopausal women, presented in Chapter 8, granted an opportunity to directly compare the role of the different pathways between these cancers because the same biomarkers were measured for the three cancers using the same assays. Although, as previously mentioned, there was high uncertainty around the estimated mediated effects, results from this study hinted that the pathways might indeed have different importance between the three cancers. This was most evident for estradiol, which explained the effect of adiposity on ER-positive breast and endometrial cancer risk in part but did not appear to have any mediating effect for colorectal cancer risk. Fasting insulin had a strong mediating effect for ER-positive breast cancer, while there was some suggestion, albeit with high uncertainty, that it might have also had a weak to moderate mediating role in the effects of adiposity on colorectal and endometrial cancer. Similarly, the point estimates indicated a likely mediating role for the inflammatory status pathway for the three cancers, but due to high uncertainty around the estimated indirect effects, it was difficult to compare the importance of this pathway between the cancers.

9.2 Strengths and limitations

Strengths and limitations of the individual studies have been discussed in detail in the relevant chapters. This section provides an overview of the strengths and limitations of the thesis.

This thesis made use of existing data from four prospective studies, three of which (UK Biobank, EPIC, and WHI-OS) had at least one biomarker measured to reflect each of the three pathways hypothesised to link adiposity to cancer, and one (MCCS) had biomarkers

reflecting two of the pathways. These studies had rich data on potential confounders and in all, except for a subgroup in the EPIC cohort (see Table 4-1), anthropometric data were collected by direct measurement. The causal mediation analysis methods applied to analyse the data allowed accounting for potential dependencies between the assessed biomarkers. Access to these data and the statistical analysis approaches used were the major strengths of this thesis. Collectively these cohorts and causal mediation methods enabled a comprehensive investigation of the role of inflammatory status, insulin resistance and hyperinsulinemia, and sex-steroid hormones in explaining the effect of adiposity on the risk of colorectal cancer in men and postmenopausal women, and ER-positive breast and endometrial cancers in postmenopausal women. However, the estimated mediated effects for each pathway must be interpreted in light of several issues that are explained below.

An observed mediating role through a given biomarker does not necessarily imply that the biomarker has a causal effect on cancer risk or that a direct intervention on the biomarker would influence cancer risk. For example, in the mediation analysis of the EPIC data, in Chapter 7, of the assessed biomarkers, pathways originating with reduced adiponectin level in the obese state had the largest mediating effect on the adiposity-endometrial cancer risk association. This analysis was sequential, assuming that adiponectin was the most upstream biomarker, i.e. it was assumed that adiponectin could have influenced other biomarkers, but not vice versa. Even if this assumption holds, it would, still, be naïve to deduce from this observation that intervening to increase the adiponectin level of women in the obese state, supposing the intervention existed, would necessarily have an effect in reducing their endometrial cancer risk. The estimated indirect effect through adiponectin might have, at least in part, been because measured

adiponectin performed well at capturing the inflammatory status in the obese state, which in turn could have influenced cancer risk through several mechanisms.

Following the above point, an estimated mediating role through a given pathway would have been influenced by how well the biomarkers reflected the physiological impact of that pathway. For example, in the UK Biobank study (Chapter 5), CRP, a non-specific biomarker of systemic inflammation, was the only inflammatory biomarker available and it is unlikely that it would have reflected all features of the inflammatory status pathway. Similarly, in the same study, the available biomarkers representing the sex-steroid hormone pathway, i.e. SHBG and testosterone, would have only captured some aspects of dysregulation in this pathway in the obese state.

The estimated mediating effects through the pathways would have also been affected by the quality of measurements. In Table 4-3, the total, within, and between assay coefficients of variation (CV) for the measured biomarkers in the four studies were provided, suggesting that in all studies biomarker measurements were in general of good quality. Nonetheless, there were differences in the measurement quality of the biomarkers, which might have influenced the estimated effects to varying degrees. Additionally, for a given pathway captured by several biomarkers (for example in the EPIC study in Chapter 7, measures for adiponectin, IL-6, IL-1Ra, TNF-R1, TNF-R2, and CRP were used together to capture the inflammatory status) accumulation of errors might have made further contributions to imprecision around the estimated indirect effect through that pathway. Finally, in all studies, biomarker measurements at a single time point were used, and differences in temporal variability might again have influenced the estimated effects to varying degrees.

Access to several datasets provided an opportunity to explore the adiposity-cancer link for three cancer sites. Also, for postmenopausal women, for each cancer site two datasets were analysed (UK Biobank and WHI-OS for colorectal cancer, MCCS and WHI-OS for ER-positive breast cancer, and EPIC and WHI-OS for endometrial cancer). However, the results of these studies are not necessarily directly comparable. Firstly, the studies have different designs and populations and across the studies, the biomarkers used to capture the pathways were either not the same or were measured using different assays (see Table 4-3). Additionally, different designs of the studies called for different approaches to mediation analysis. For example, for colorectal cancer, a sequential mediation analysis was used to estimate the natural indirect (NIE) and direct (NDE) effects in the UK Biobank study. In contrast, an interventional analysis was used to estimate the interventional indirect (IIE) and direct (IDE) effects in the WHI-OS study. Similarly, for endometrial cancer, in the EPIC study, the NIEs and NDE were estimated using the sequential approach, but in the WHI-OS, the IIEs and IDE were estimated. As illustrated in Figure 3-10 (for the sequential NIEs and NDE) and Figure 3-13 (for the IIEs and IDE), the interpretations of these effects are different. Briefly, if there are three causally ordered mediators $M_1 \rightarrow M_2 \rightarrow M_3$ linking A (exposure) to Y (outcome), using the sequential analysis, the estimated NIE through M_1 captures the effects through all pathways originating with M_1, i.e. $A \rightarrow M_1 \rightarrow Y$, $A \rightarrow M_1 \rightarrow M_2 \rightarrow Y$, and $A \rightarrow M_1 \rightarrow M_2 \rightarrow M_3 \rightarrow Y$ while the NIE through M_3 captures the indirect effect through M_3, beyond the potential influences of M_1 and M_2 on levels of M_3, i.e. $A \rightarrow M_3 \rightarrow Y$. Conversely, using the interventional approach, the IIE through M_1 captures the indirect effect through M_1, but not through its causal descendants, i.e. $A \rightarrow M_1 \rightarrow Y$, while the IIE though M_3 captures the effects $A \rightarrow M_3 \rightarrow Y$, $A \rightarrow M_2 \rightarrow M_3 \rightarrow Y$, and $A \rightarrow M_1 \rightarrow M_2 \rightarrow M_3 \rightarrow Y$.

Another limitation of this research was that, in all studies, measures for adiposity and biomarkers were collected cross-sectionally at baseline. Therefore, reverse causation for adiposity-biomarker associations cannot be ruled out.

9.3 Discussion and implications of the findings

Adiposity and colorectal cancer risk in men and postmenopausal women

Existing studies investigating the adiposity-colorectal cancer link (59, 83, 117, 133, 134) (summarised in Table 3-2) indicate that alterations in levels of inflammatory biomarkers might partly explain increased cancer risk in men and women with obesity. However, previous studies, as well as the analysis of the UK Biobank data presented in Chapter 5 have only shown small mediating effects through CRP, suggesting that a broader range of adipocytokines and cytokines might be needed to better characterise the obesity-induced inflammatory status. In the analysis of the WHI-OS case-cohort study of postmenopausal women, presented in Chapter 8, several biomarkers were used in combination to capture the inflammatory status pathway, and the point estimate indicated a larger mediating effect through this pathway. There was, however, high uncertainty around the estimated indirect effect. Also, previously, in the Health Professionals Follow-Up Study (HPFS) of men, using a causal mediation analysis approach that took dependencies between biomarkers into account, strong mediation was observed through CRP, IL-6, TNFR-2, macrophage inhibitory cytokine-1 (MIC-1), adiponectin, soluble leptin receptor (sOB-R), and C-peptide combined, even though the indirect effects through the individual biomarkers were mostly small (Table 3-2) (133).

In Chapter 5, no mediating effect through HbA1c (a proxy biomarker for insulin resistance) beyond the influence of CRP on adiposity-colorectal cancer risk association

was observed in men or postmenopausal women. Similarly, previously in a case-control study nested within the EPIC, adjusting for HbA1c almost made no difference to the association between waist circumference and colon cancer in men or women (117). From these findings, it could be conjectured that on its own, HbA1c might not be a suitable biomarker of a dysregulated insulin signalling pathway. In another causal mediation analysis, a mediating role, although weak, was observed for triglyceride-glucose index (TyG), which again was used as a proxy for insulin resistance (134) (Table 3-2). In Chapter 8, a weak to moderate mediating effect through fasting insulin was observed for postmenopausal women in the WHI-OS case-cohort study, but there was high uncertainty around the estimated effect. Previous studies within the same cohort have also supported a possible mediating role for fasting insulin in postmenopausal women (59) (83). Also, there has been a suggestion for a possible mediating role for C-peptide in women (117), but not in men (117, 133) (Table 3-2). It is difficult to comment on the presence and importance of a mediating role of the insulin signalling pathway in the effect of adiposity on colorectal cancer risk based on these observations and additional studies with larger sample sizes are needed to investigate the role of this pathway further.

Studies on the role of the sex-steroid hormone pathway in explaining the effect of adiposity on colorectal cancer risk are lacking. In Chapter 5 of this thesis, SHBG and testosterone levels were used jointly to capture some characteristics of a dysregulated sex-steroid hormone pathway. In men with obesity, both of these biomarkers were reduced and explained a small part of the adiposity-colorectal cancer risk association. In postmenopausal women with obesity, SHBG levels were reduced, and testosterone levels increased, but these alterations did not appear to have any mediating effect in the association. In the analysis of the WHI-OS data of postmenopausal women, presented in

Chapter 8, estradiol was the only available biomarker on the sex-steroid hormone pathway, through which, again an effectively null indirect effect was observed. The mediating effect through reduced SHBG and testosterone levels observed for men in the UK Biobank study is supported by the existing evidence for the potential protective role of androgens in colorectal cancer carcinogenesis (18). If replicated, the observed mediating role through SHBG and testosterone in men in the UK Biobank study, coupled with the null mediating effects through the sex-steroid hormone pathway in postmenopausal women seen in the UK Biobank and WHI-OS studies, suggest that rather than the increased endogenous estrogens in women, which has been hypothesised to contribute to the weaker BMI-colorectal cancer association and lower colorectal cancer incidence observed in women compared with men (8, 9), these differences might, in part, be explained by reduced androgen levels in men with obesity.

Adiposity and breast cancer risk in postmenopausal women

In Chapter 8, analysis of the WHI-OS case-cohort study showed a weak to moderate mediating role for total estradiol in adiposity-ER-positive postmenopausal breast cancer risk association. A moderate to strong mediating role for calculated free estradiol was also observed in the MCCS case-cohort study, presented in Chapter 5. These observations are in line with current evidence (summarised in Table 3-3), suggesting that estradiol, especially free estradiol, plays a major mediating role in explaining the effect of adiposity on postmenopausal breast cancer risk (16, 68, 119-121, 135). However, they also indicate that despite its importance, [free] estradiol does not explain all of the effect of adiposity on postmenopausal breast cancer risk.

In a previous causal mediation analysis of two case-cohort studies within the WHI-OS (120), one of which was the same study used in Chapter 8 of this thesis, fasting insulin showed a strong mediating role in the effect of adiposity on postmenopausal breast cancer risk (Table 3-3). This strong mediating effect through fasting insulin was replicated in the analysis of the WHI-OS case-cohort study presented in Chapter 8, which contrary to the previous study (120), also took the potential confounding effect of inflammatory status on the hyperinsulinemia-breast cancer association into account. In the MCCS study, in Chapter 6, only weak mediation was observed for fasting insulin, with high uncertainty around the estimated indirect effects. A variety of factors, such as sample variability, different assays used to measure insulin, and different distribution of confounders, might have contributed to the different mediating roles observed for fasting insulin in these two investigations.

Studies of the role of inflammatory status in the effect of adiposity on postmenopausal breast cancer risk are limited. One case-cohort study within the WHI-OS demonstrated attenuation in the hazard ratio after adjusting for CRP (Table 3-3) (68). This thesis also had limited capacity in exploring the role of this pathway, because of the studies on breast cancer, only one of the cohorts (WHI-OS) had measures for inflammatory biomarkers. This study, presented in Chapter 8, yielded inconsistent and imprecise estimates for an indirect effect through inflammatory status pathway for different measures of adiposity. Additional studies with larger sample sizes and measurements for a range of biomarkers reflecting the inflammatory status are needed to investigate the role of this pathway in the effect of adiposity on postmenopausal breast cancer risk.

Adiposity and endometrial cancer risk in postmenopausal women

The predominant theory for explaining the increased endometrial cancer risk in postmenopausal women with obesity is enhanced production of estrogens in adipose tissue (112, 190). In previous studies (summarised in Table 3-4), adjusting for estradiol or estrone attenuated the adiposity-endometrial cancer association only in part (76, 77, 114). In Chapter 7, analysis of the case-control study nested within the EPIC showed weak evidence for a small to moderate mediating effect through estrogens beyond the potential influences of inflammatory status and C-peptide. In the WHI-OS case-cohort study, presented in Chapter 8, moderate mediation was observed through estradiol. These observations provide additional evidence for the role of estrogens in explaining the effect of adiposity on endometrial cancer development in postmenopausal women, but also indicate that other mechanisms are also likely to be involved.

In both studies of endometrial cancer in this thesis, there were suggestions for a mediating effect through the inflammatory status pathway, with point estimates indicating a larger mediation through this pathway compared with the sex-steroid hormone pathway. However, in the WHI-OS (Chapter 8), the evidence was only weak, and there was high uncertainty around the estimated indirect effect through this pathway. In line with these results, in previous studies an attenuation in the adiposity-endometrial cancer association was reported after adjusting for inflammatory biomarkers (e.g. CRP, IL-6, or IL-1Ra) (76, 77) (see Table 3-4). The potential role of inflammation in endometrial cancer carcinogenesis in the obese state is further supported by observational studies reporting an inverse association between use of nonsteroidal anti-inflammatory drugs including aspirin and endometrial cancer risk in women with obesity (71, 72).

In previous studies of adiposity and endometrial cancer, adjusting for fasting insulin or C-peptide attenuated the association to some extent (76, 77, 94), while in a causal mediation analysis, the estimated indirect effect for TyG (used as a proxy measure for insulin resistance) on BMI-endometrial cancer association was essentially null (134) (Table 3-4). In the studies of endometrial cancer included in this thesis, there were suggestions for weak mediation through C-peptide, beyond the potential influence of inflammatory status pathway (Chapter 7), or weak to moderate mediation fasting insulin (Chapter 8), but there were high uncertainties around the estimated indirect effects. It is, therefore, not possible to comment on the role and importance of this pathway in adiposity-endometrial cancer association based on current evidence.

9.4 Future research

Results from this research need to be replicated in studies with larger sample sizes, a wide range of biomarkers reflecting the inflammatory status, insulin resistance and hyperinsulinemia, and alterations in sex-steroid hormone levels, and ideally repeated measurements for adiposity and biomarkers. Larger studies are needed to not only produce results with more certainty, but also to allow exploring the possible influences of adiposity-biomarker and biomarker-biomarker interactions on observed mediating effects through the biomarkers. Ideally, future studies would take advantage of novel causal mediation analysis methods to appropriately account for dependencies between biomarkers.

Correct characterisation of the pathways based on biomarkers plays a crucial role in advancing knowledge about mechanisms underlying the effect of adiposity on cancer risk. This is especially challenging for more complex pathways, such as inflammatory

status, which is likely defined based on an intricate network of cellular, molecular, and metabolic mechanisms. While it is appealing to measure and combine a wide range of biomarkers in an attempt to capture a pathway in all its complexities, increasing the number of biomarkers might come at the expense of accumulating measurement errors, thereby reducing precision around estimated effects. Also, developing methods for correcting direct and indirect effects for measurement error in the context of mediation analysis with multiple mediators, which are currently lacking, would improve the identification and quantification of mechanistic pathways. Additionally, while it might be viable to combine several biomarkers without using any data reduction approach, modelling a larger number of biomarkers, for example as produced by metabolomic studies, would get complicated and even impossible. Identifying ideal data reduction strategies for mediation analysis is another area of research that could make considerable contributions to future studies on mechanistic pathways linking adiposity to cancer risk.

Mediation analysis relies on strong no unmeasured confounding assumptions. Extending sensitivity analysis and bias correction approaches that currently exist for simple scenarios with single mediators (see for example references (200, 201)) for settings with multiple mediators could be important for investigating the potential influences of unmeasured confounding on estimated indirect and direct effects.

In the field of biomarkers and cancer epidemiology, three study designs, including cohort, case-cohort, and nested case-control studies, are dominant. In this thesis, for analysis of the cohort study (UK Biobank, Chapter 5), cancer diagnosis was treated as the binary outcome and participants lost to follow-up were excluded from the analytic dataset. Also, the two case-cohort studies (MCCS, Chapter 6 and WHI-OS Chapter 8) were treated as

case-control studies with cumulative sampling. These choices were made because methods for mediation analysis to handle multiple mediators and survival outcome are not fully developed yet. Developing these methods could be of great use for investigating the role of biomarkers in connecting adiposity (and other risk factors) to cancer incidence. For the nested case-control study design, sequential mediation analysis using the regression-based method can be modified to accommodate the data (see Table 3-1). It would be helpful if future work also extends the interventional mediation analysis using regression-standardisation method (130) for nested case-control data. As previously discussed, (see **3.7.3 Interventional analysis**), interventional effects can be used without needing to assume a causal ordering of mediators, which compared with the sequential mediation analysis is a considerable advantage.

9.5 Conclusions

This thesis provided an investigation of the mediating roles of inflammatory status, insulin resistance and hyperinsulinemia, and altered sex-steroid hormone levels in explaining the effect of adiposity on the risk of colorectal cancer in men and postmenopausal women, and ER-positive breast and endometrial cancer in postmenopausal women. For colorectal cancer, it offered additional evidence for the role of inflammatory status in explaining the effect in men and postmenopausal women but did not add to a scarcity of evidence for the mediating role of insulin signalling pathway. It also found no evidence of mediation through sex-steroid hormones for women, but a small mediating effect through reduced SHBG and testosterone levels for men. For ER-positive breast cancer in postmenopausal women, it added to the existing body of evidence for the important mediating role of estradiol, while showing that it did not explain all of the effect of adiposity. It also suggested that fasting insulin played a

mediating role, but it made little contribution to the already lacking evidence for a role through inflammatory status. For endometrial cancer in postmenopausal women, this research demonstrated that although estrogens played a mediating role, they did not explain all of the effect, and that inflammatory status also had a potential mediating role. It, however, did not show evidence for a mediating role through insulin signalling pathway.

Studies with larger sample sizes and repeated measurements for adiposity and biomarkers are needed to explore further the mediating role of these pathways in explaining the effect of adiposity on cancer risk. Causal mediation analyses methods that allow investigating and disentangling the role of multiple dependent mediators need to be taken advantage of by these future studies. Methodological advances in mediation analysis, including developing methods that can handle survival outcome or accommodate data from nested case-control studies, could make considerable contributions to this field. Accurate characterisation of the involved pathways based on measured biomarkers is a challenging, yet crucial task for advancing our knowledge of the role of different biological pathways in explaining the effect of adiposity on cancer risk. Novel technical advances already provide access to a wide range of biomarker measurements. Future research would benefit from exploring approaches that allow exploiting these data in the context of mediation analysis.

References

1. IARC (2018). Absence of excess body fatness. IARC Handb Cancer Prev. 16:1–646. Available from: http://publications.iarc.fr/570.
2. World Cancer Research Fund/American Institute for Cancer Research. Diet, Nutrition, Physical Activity and Cancer: a Global Perspective. Continuous Update Project Expert Rerpor 2018. Available at http://dietandcancerreport.org/.
3. Arnold M, Pandeya N, Byrnes G, Renehan PAG, Stevens GA, Ezzati PM, et al. Global burden of cancer attributable to high body-mass index in 2012: a population-based study. Lancet Oncol. 2015;16(1):36-46.
4. N. C. D. Risk Factor Collaboration. Worldwide trends in body-mass index, underweight, overweight, and obesity from 1975 to 2016: a pooled analysis of 2416 population-based measurement studies in 128.9 million children, adolescents, and adults. Lancet. 2017;390(10113):2627-42.
5. Hanahan D, Weinberg RA. Hallmarks of cancer: the next generation. Cell. 2011;144(5):646-74.
6. Deng T, Lyon CJ, Bergin S, Caligiuri MA, Hsueh WA. Obesity, Inflammation, and Cancer. Annu Rev Pathol. 2016;11:421-49.
7. Gregor MF, Hotamisligil GS. Inflammatory mechanisms in obesity. Annu Rev Immunol. 2011;29:415-45.
8. Hardy OT, Czech MP, Corvera S. What causes the insulin resistance underlying obesity? Curr Opin Endocrinol Diabetes Obes. 2012;19(2):81-7.
9. Yadav A, Kataria MA, Saini V, Yadav A. Role of leptin and adiponectin in insulin resistance. Clin Chim Acta. 2013;417:80-4.
10. Gerard C, Brown KA. Obesity and breast cancer - Role of estrogens and the molecular underpinnings of aromatase regulation in breast adipose tissue. Mol Cell Endocrinol. 2017;466:15-30.
11. Grossmann M. Hypogonadism and male obesity: Focus on unresolved questions. Clin Endocrinol (Oxf). 2018;89(1):11-21.
12. Simo R, Saez-Lopez C, Barbosa-Desongles A, Hernandez C, Selva DM. Novel insights in SHBG regulation and clinical implications. Trends Endocrinol Metab. 2015;26(7):376-83.
13. Nimptsch K, Pischon T. Obesity Biomarkers, Metabolism and Risk of Cancer: An Epidemiological Perspective. Recent Results Cancer Res. 2016;208:199-217.
14. Renehan AG, Frystyk J, Flyvbjerg A. Obesity and cancer risk: the role of the insulin-IGF axis. Trends Endocrinol Metab. 2006;17(8):328-36.
15. Pollak M. Insulin and insulin-like growth factor signalling in neoplasia. Nat Rev Cancer. 2008;8(12):915-28.
16. Key TJ, Appleby PN, Reeves GK, Roddam A, Dorgan JF, Longcope C, et al. Body mass index, serum sex hormones, and breast cancer risk in postmenopausal women. J Natl Cancer Inst. 2003;95(16):1218-26.
17. Kapoor D, Malkin CJ, Channer KS, Jones TH. Androgens, insulin resistance and vascular disease in men. Clin Endocrinol (Oxf). 2005;63(3):239-50.
18. Lin JH, Giovannucci E. Sex hormones and colorectal cancer: what have we learned so far? J Natl Cancer Inst. 2010;102(23):1746-7.
19. VanderWeele T. Explanation in causal inference: methods for mediation and interaction. New York: Oxford Univ. Press; 2015.
20. Varmus H, Harlow E. Science funding: Provocative questions in cancer research. Nature. 2012;481(7382):436-7.
21. National Health and Medical Research Council. Clinical practice guidelines for the management of overweight and obesity in adults, adolescents and children in Australia. Melbourne: National Health and Medical Research Council, 2013. .

22. World Health Organization. Obesity and Overweight Fact Sheet. World Health Organization, Geneva. Available from: http://www.who.int/mediacentre/factsheets/fs311/en/ , accessed on 10/07/2017. 2016.
23. IARC Working Group on the Evaluation of Cancer-Preventive Interventions. Absence of excess body fatness, IARC Handbooks of Cancer Prevention ; Volume 16. 2018.
24. World Cancer Research Fund/American Institute for Cancer Research. Continuous Update Project Expect Report 2018. Body fatness and weight gain and the risk of cancer. Available at http://dietandcancerreport.org/.
25. World Health Organization. Waist Circumference and Waist-Hip Ratio: Report of a WHO Expert Consultation. https://apps.who.int/iris/bitstream/handle/10665/44583/9789241501491_eng.pdf?sequence=1; 2008.
26. International Diabetes Federation. The IDF consensus worldwide definition of the metabolic syndorme 2006. Available at https://www.idf.org
27. International Diabetes Federation. The IDF consensus worldwide definition of the metabolic syndorme 2006. Available at https://www.idf.org
28. N. C. D. Risk Factor Collaboration. Trends in adult body-mass index in 200 countries from 1975 to 2014: a pooled analysis of 1698 population-based measurement studies with 19.2 million participants. Lancet. 2016;387(10026):1377-96.
29. World Cancer Research Fund/American Institute for Cancer Research. Diet, Nutrition, Physical Activity and Cancer: a Global Perspective. Continuous Update Project Expert Rerpor 2018. Available at http://dietandcancerreport.org/.
30. Lauby-Secretan B, Scoccianti C, Loomis D, Grosse Y, Bianchini F, Straif K, et al. Body Fatness and Cancer--Viewpoint of the IARC Working Group. N Engl J Med. 2016;375(8):794-8.
31. Bray F, Ferlay J, Soerjomataram I, Siegel RL, Torre LA, Jemal A. Global cancer statistics 2018: GLOBOCAN estimates of incidence and mortality worldwide for 36 cancers in 185 countries. CA Cancer J Clin. 2018;68(6):394-424.
32. Arnold M, Sierra MS, Laversanne M, Soerjomataram I, Jemal A, Bray F. Global patterns and trends in colorectal cancer incidence and mortality. Gut. 2017;66(4):683-91.
33. Ma Y, Yang Y, Wang F, Zhang P, Shi C, Zou Y, et al. Obesity and risk of colorectal cancer: a systematic review of prospective studies. PLoS One. 2013;8(1):e53916.
34. World Cancer Research Fund International/American Institute for Cancer Research. Continuous Update Project Report: Diet, Nutrition, Physical Activity and Colorectal Cancer. Available at: wcrf.org/colorectal-cancer-2017; 2017.
35. Cornish AJ, Law PJ, Timofeeva M, Palin K, Farrington SM, Palles C, et al. Modifiable pathways for colorectal cancer: a mendelian randomisation analysis. Lancet Gastroenterol Hepatol. 2019.
36. Gao C, Patel CJ, Michailidou K, Peters U, Gong J, Schildkraut J, et al. Mendelian randomization study of adiposity-related traits and risk of breast, ovarian, prostate, lung and colorectal cancer. Int J Epidemiol. 2016;45(3):896-908.
37. Thrift AP, Gong J, Peters U, Chang-Claude J, Rudolph A, Slattery ML, et al. Mendelian Randomization Study of Body Mass Index and Colorectal Cancer Risk. Cancer Epidemiol Biomarkers Prev. 2015;24(7):1024-31.
38. Ginsburg O, Bray F, Coleman MP, Vanderpuye V, Eniu A, Kotha SR, et al. The global burden of women's cancers: a grand challenge in global health. Lancet. 2017;389(10071):847-60.
39. Torre LA, Islami F, Siegel RL, Ward EM, Jemal A. Global Cancer in Women: Burden and Trends. Cancer Epidemiol Biomarkers Prev. 2017;26(4):444-57.
40. Munsell MF, Sprague BL, Berry DA, Chisholm G, Trentham-Dietz A. Body mass index and breast cancer risk according to postmenopausal estrogen-progestin use and hormone receptor status. Epidemiol Rev. 2014;36:114-36.

41. World Cancer Research Fund International/American Institute for Cancer Research. Continuous Update Project Report: Diet, Nutrition, Physical Activity and Breast Cancer. Available at: wcrf.org/breast-cancer-2017; 2017.
42. Guo Y, Warren Andersen S, Shu XO, Michailidou K, Bolla MK, Wang Q, 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(8):e1002105.
43. Shu X, Wu L, Khankari NK, Shu XO, Wang TJ, Michailidou K, et al. Associations of obesity and circulating insulin and glucose with breast cancer risk: a Mendelian randomization analysis. Int J Epidemiol. 2019;48(3):795-806.
44. Lortet-Tieulent J, Ferlay J, Bray F, Jemal A. International Patterns and Trends in Endometrial Cancer Incidence, 1978-2013. J Natl Cancer Inst. 2018;110(4):354-61.
45. World Cancer Research Fund International/American Institute for Cancer Research. Continuous Update Project Report: Diet, Nutrition, Physical Activity and endometrial cancer. Available at: https://www.wcrf.org/dietandcancer/contents; 2018.
46. Jenabi E, Poorolajal J. The effect of body mass index on endometrial cancer: a meta-analysis. Public Health. 2015;129(7):872-80.
47. World Cancer Research Fund International/American Institute for Cancer Research. WCRF/AICR Systematic Literature Review Continuous Update Project Report: The Associations between Food, Nutrition and Physical Activity and the Risk of Endometrial Cancer. 2012.
48. Setiawan VW, Yang HP, Pike MC, McCann SE, Yu H, Xiang YB, et al. Type I and II endometrial cancers: have they different risk factors? J Clin Oncol. 2013;31(20):2607-18.
49. Nead KT, Sharp SJ, Thompson DJ, Painter JN, Savage DB, Semple RK, et al. Evidence of a Causal Association Between Insulinemia and Endometrial Cancer: A Mendelian Randomization Analysis. J Natl Cancer Inst. 2015;107(9).
50. Painter JN, O'Mara TA, Marquart L, Webb PM, Attia J, Medland SE, et al. Genetic Risk Score Mendelian Randomization Shows that Obesity Measured as Body Mass Index, but not Waist:Hip Ratio, Is Causal for Endometrial Cancer. Cancer Epidemiol Biomarkers Prev. 2016;25(11):1503-10.
51. Amitani M, Asakawa A, Amitani H, Inui A. The role of leptin in the control of insulin-glucose axis. Front Neurosci. 2013;7:51.
52. Nigro E, Scudiero O, Monaco ML, Palmieri A, Mazzarella G, Costagliola C, et al. New insight into adiponectin role in obesity and obesity-related diseases. Biomed Res Int. 2014;2014:658913.
53. Chan PC, Hsiao FC, Chang HM, Wabitsch M, Hsieh PS. Importance of adipocyte cyclooxygenase-2 and prostaglandin E2-prostaglandin E receptor 3 signaling in the development of obesity-induced adipose tissue inflammation and insulin resistance. FASEB J. 2016;30(6):2282-97.
54. Gerard C, Brown KA. Obesity and breast cancer - Role of estrogens and the molecular underpinnings of aromatase regulation in breast adipose tissue. Mol Cell Endocrinol. 2017.
55. Laukoetter MG, Mennigen R, Hannig CM, Osada N, Rijcken E, Vowinkel T, et al. Intestinal cancer risk in Crohn's disease: a meta-analysis. J Gastrointest Surg. 2011;15(4):576-83.
56. Jess T, Rungoe C, Peyrin-Biroulet L. Risk of colorectal cancer in patients with ulcerative colitis: a meta-analysis of population-based cohort studies. Clin Gastroenterol Hepatol. 2012;10(6):639-45.
57. Patrignani P, Patrono C. Aspirin and Cancer. J Am Coll Cardiol. 2016;68(9):967-76.
58. Wei EK, Giovannucci E, Fuchs CS, Willett WC, Mantzoros CS. Low plasma adiponectin levels and risk of colorectal cancer in men: a prospective study. J Natl Cancer Inst. 2005;97(22):1688-94.

59. Ho GY, Wang T, Gunter MJ, Strickler HD, Cushman M, Kaplan RC, et al. Adipokines linking obesity with colorectal cancer risk in postmenopausal women. Cancer Res. 2012;72(12):3029-37.
60. Aleksandrova K, Boeing H, Jenab M, Bueno-de-Mesquita HB, Jansen E, van Duijnhoven FJ, et al. Total and high-molecular weight adiponectin and risk of colorectal cancer: the European Prospective Investigation into Cancer and Nutrition Study. Carcinogenesis. 2012;33(6):1211-8.
61. Song M, Zhang X, Wu K, Ogino S, Fuchs CS, Giovannucci EL, et al. Plasma adiponectin and soluble leptin receptor and risk of colorectal cancer: a prospective study. Cancer Prev Res (Phila). 2013;6(9):875-85.
62. Nimptsch K, Song M, Aleksandrova K, Katsoulis M, Freisling H, Jenab M, et al. Genetic variation in the ADIPOQ gene, adiponectin concentrations and risk of colorectal cancer: a Mendelian Randomization analysis using data from three large cohort studies. Eur J Epidemiol. 2017;32(5):419-30.
63. Tamakoshi K, Toyoshima H, Wakai K, Kojima M, Suzuki K, Watanabe Y, et al. Leptin is associated with an increased female colorectal cancer risk: a nested case-control study in Japan. Oncology. 2005;68(4-6):454-61.
64. Aleksandrova K, Boeing H, Jenab M, Bueno-de-Mesquita HB, Jansen E, van Duijnhoven FJ, et al. Leptin and soluble leptin receptor in risk of colorectal cancer in the European Prospective Investigation into Cancer and Nutrition cohort. Cancer Res. 2012;72(20):5328-37.
65. Zhou B, Shu B, Yang J, Liu J, Xi T, Xing Y. C-reactive protein, interleukin-6 and the risk of colorectal cancer: a meta-analysis. Cancer Causes Control. 2014;25(10):1397-405.
66. Wang X, Dai JY, Albanes D, Arndt V, Berndt SI, Bezieau S, et al. Mendelian randomization analysis of C-reactive protein on colorectal cancer risk. Int J Epidemiol. 2019;48(3):767-80.
67. Macis D, Guerrieri-Gonzaga A, Gandini S. Circulating adiponectin and breast cancer risk: a systematic review and meta-analysis. Int J Epidemiol. 2014;43(4):1226-36.
68. Gunter MJ, Wang T, Cushman M, Xue X, Wassertheil-Smoller S, Strickler HD, et al. Circulating Adipokines and Inflammatory Markers and Postmenopausal Breast Cancer Risk. J Natl Cancer Inst. 2015;107(9).
69. Chan DS, Bandera EV, Greenwood DC, Norat T. Circulating C-Reactive Protein and Breast Cancer Risk-Systematic Literature Review and Meta-analysis of Prospective Cohort Studies. Cancer Epidemiol Biomarkers Prev. 2015;24(10):1439-49.
70. Robinson T, Martin RM, Yarmolinsky J. Mendelian Randomization Analysis of Circulating Adipokines and C-reactive Protein on Breast Cancer Risk. bioRxiv. 2019:720110.
71. Verdoodt F, Friis S, Dehlendorff C, Albieri V, Kjaer SK. Non-steroidal anti-inflammatory drug use and risk of endometrial cancer: A systematic review and meta-analysis of observational studies. Gynecol Oncol. 2016;140(2):352-8.
72. Webb PM, Na R, Weiderpass E, Adami HO, Anderson KE, Bertrand KA, et al. Use of aspirin, other nonsteroidal anti-inflammatory drugs and acetaminophen and risk of endometrial cancer: the Epidemiology of Endometrial Cancer Consortium. Ann Oncol. 2019;30(2):310-6.
73. Zeng F, Shi J, Long Y, Tian H, Li X, Zhao AZ, et al. Adiponectin and Endometrial Cancer: A Systematic Review and Meta-Analysis. Cell Physiol Biochem. 2015;36(4):1670-8.
74. Luhn P, Dallal CM, Weiss JM, Black A, Huang WY, Lacey JV, Jr., et al. Circulating adipokine levels and endometrial cancer risk in the prostate, lung, colorectal, and ovarian cancer screening trial. Cancer Epidemiol Biomarkers Prev. 2013;22(7):1304-12.
75. Dallal CM, Brinton LA, Bauer DC, Buist DS, Cauley JA, Hue TF, et al. Obesity-related hormones and endometrial cancer among postmenopausal women: a nested case-control study within the B~FIT cohort. Endocr Relat Cancer. 2013;20(1):151-60.

76. Wang T, Rohan TE, Gunter MJ, Xue X, Wactawski-Wende J, Rajpathak SN, et al. A prospective study of inflammation markers and endometrial cancer risk in postmenopausal hormone nonusers. Cancer Epidemiol Biomarkers Prev. 2011;20(5):971-7.
77. Dossus L, Rinaldi S, Becker S, Lukanova A, Tjonneland A, Olsen A, et al. Obesity, inflammatory markers, and endometrial cancer risk: a prospective case-control study. Endocr Relat Cancer. 2010;17(4):1007-19.
78. Dossus L, Becker S, Rinaldi S, Lukanova A, Tjonneland A, Olsen A, et al. Tumor necrosis factor (TNF)-alpha, soluble TNF receptors and endometrial cancer risk: the EPIC study. Int J Cancer. 2011;129(8):2032-7.
79. Petersen MC, Shulman GI. Mechanisms of Insulin Action and Insulin Resistance. Physiol Rev. 2018;98(4):2133-223.
80. Shimobayashi M, Albert V, Woelnerhanssen B, Frei IC, Weissenberger D, Meyer-Gerspach AC, et al. Insulin resistance causes inflammation in adipose tissue. J Clin Invest. 2018;128(4):1538-50.
81. Tsilidis KK, Kasimis JC, Lopez DS, Ntzani EE, Ioannidis JP. Type 2 diabetes and cancer: umbrella review of meta-analyses of observational studies. BMJ (Clinical research ed). 2015;350:g7607.
82. Sherwani SI, Khan HA, Ekhzaimy A, Masood A, Sakharkar MK. Significance of HbA1c Test in Diagnosis and Prognosis of Diabetic Patients. Biomark Insights. 2016;11:95-104.
83. Gunter MJ, Hoover DR, Yu H, Wassertheil-Smoller S, Rohan TE, Manson JE, et al. Insulin, insulin-like growth factor-I, endogenous estradiol, and risk of colorectal cancer in postmenopausal women. Cancer Res. 2008;68(1):329-37.
84. Limburg PJ, Stolzenberg-Solomon RZ, Vierkant RA, Roberts K, Sellers TA, Taylor PR, et al. Insulin, glucose, insulin resistance, and incident colorectal cancer in male smokers. Clin Gastroenterol Hepatol. 2006;4(12):1514-21.
85. Jenab M, Riboli E, Cleveland RJ, Norat T, Rinaldi S, Nieters A, et al. Serum C-peptide, IGFBP-1 and IGFBP-2 and risk of colon and rectal cancers in the European Prospective Investigation into Cancer and Nutrition. Int J Cancer. 2007;121(2):368-76.
86. Ma J, Giovannucci E, Pollak M, Leavitt A, Tao Y, Gaziano JM, et al. A prospective study of plasma C-peptide and colorectal cancer risk in men. J Natl Cancer Inst. 2004;96(7):546-53.
87. Otani T, Iwasaki M, Sasazuki S, Inoue M, Tsugane S, Japan Public Health Center-based Prospective Study G. Plasma C-peptide, insulin-like growth factor-I, insulin-like growth factor binding proteins and risk of colorectal cancer in a nested case-control study: the Japan public health center-based prospective study. Int J Cancer. 2007;120(9):2007-12.
88. Rinaldi S, Rohrmann S, Jenab M, Biessy C, Sieri S, Palli D, et al. Glycosylated hemoglobin and risk of colorectal cancer in men and women, the European prospective investigation into cancer and nutrition. Cancer Epidemiol Biomarkers Prev. 2008;17(11):3108-15.
89. Joshu CE, Prizment AE, Dluzniewski PJ, Menke A, Folsom AR, Coresh J, et al. Glycated hemoglobin and cancer incidence and mortality in the Atherosclerosis in Communities (ARIC) Study, 1990-2006. Int J Cancer. 2012;131(7):1667-77.
90. Gunter MJ, Hoover DR, Yu H, Wassertheil-Smoller S, Manson JE, Li J, et al. A prospective evaluation of insulin and insulin-like growth factor-I as risk factors for endometrial cancer. Cancer Epidemiol Biomarkers Prev. 2008;17(4):921-9.
91. Autier P, Koechlin A, Boniol M, Mullie P, Bolli G, Rosenstock J, et al. Serum insulin and C-peptide concentration and breast cancer: a meta-analysis. Cancer Causes Control. 2013;24(5):873-83.

92. Lin J, Ridker PM, Rifai N, Lee IM, Manson JE, Buring JE, et al. A prospective study of hemoglobin A1c concentrations and risk of breast cancer in women. Cancer Res. 2006;66(5):2869-75.
93. Lukanova A, Zeleniuch-Jacquotte A, Lundin E, Micheli A, Arslan AA, Rinaldi S, et al. Prediagnostic levels of C-peptide, IGF-I, IGFBP -1, -2 and -3 and risk of endometrial cancer. Int J Cancer. 2004;108(2):262-8.
94. Cust AE, Allen NE, Rinaldi S, Dossus L, Friedenreich C, Olsen A, et al. Serum levels of C-peptide, IGFBP-1 and IGFBP-2 and endometrial cancer risk; results from the European prospective investigation into cancer and nutrition. Int J Cancer. 2007;120(12):2656-64.
95. Nishizawa H, Shimomura I, Kishida K, Maeda N, Kuriyama H, Nagaretani H, et al. Androgens decrease plasma adiponectin, an insulin-sensitizing adipocyte-derived protein. Diabetes. 2002;51(9):2734-41.
96. Page ST, Herbst KL, Amory JK, Coviello AD, Anawalt BD, Matsumoto AM, et al. Testosterone administration suppresses adiponectin levels in men. Journal of andrology. 2005;26(1):85-92.
97. Bosetti C, Bravi F, Negri E, La Vecchia C. Oral contraceptives and colorectal cancer risk: a systematic review and meta-analysis. Hum Reprod Update. 2009;15(5):489-98.
98. Rennert G, Rennert HS, Pinchev M, Lavie O, Gruber SB. Use of hormone replacement therapy and the risk of colorectal cancer. J Clin Oncol. 2009;27(27):4542-7.
99. Grodstein F, Martinez ME, Platz EA, Giovannucci E, Colditz GA, Kautzky M, et al. Postmenopausal hormone use and risk for colorectal cancer and adenoma. Annals of internal medicine. 1998;128(9):705-12.
100. Johnson JR, Lacey JV, Jr., Lazovich D, Geller MA, Schairer C, Schatzkin A, et al. Menopausal hormone therapy and risk of colorectal cancer. Cancer Epidemiol Biomarkers Prev. 2009;18(1):196-203.
101. Chlebowski RT, Wactawski-Wende J, Ritenbaugh C, Hubbell FA, Ascensao J, Rodabough RJ, et al. Estrogen plus progestin and colorectal cancer in postmenopausal women. N Engl J Med. 2004;350(10):991-1004.
102. Rossouw JE, Anderson GL, Prentice RL, LaCroix AZ, Kooperberg C, Stefanick ML, et al. Risks and benefits of estrogen plus progestin in healthy postmenopausal women: principal results From the Women's Health Initiative randomized controlled trial. Jama-J Am Med Assoc. 2002;288(3):321-33.
103. Simon MS, Chlebowski RT, Wactawski-Wende J, Johnson KC, Muskovitz A, Kato I, et al. Estrogen plus progestin and colorectal cancer incidence and mortality. J Clin Oncol. 2012;30(32):3983-90.
104. Rennert G. Reproductive factors, hormones and colorectal cancer-still unresolved. Br J Cancer. 2017;116(1):1-3.
105. Murphy N, Xu L, Zervoudakis A, Xue X, Kabat G, Rohan TE, et al. Reproductive and menstrual factors and colorectal cancer incidence in the Women's Health Initiative Observational Study. Br J Cancer. 2017;116(1):117-25.
106. Neumeyer S, Banbury BL, Arndt V, Berndt SI, Bezieau S, Bien SA, et al. Mendelian randomisation study of age at menarche and age at menopause and the risk of colorectal cancer. Br J Cancer. 2018.
107. Clendenen TV, Koenig KL, Shore RE, Levitz M, Arslan AA, Zeleniuch-Jacquotte A. Postmenopausal levels of endogenous sex hormones and risk of colorectal cancer. Cancer Epidemiol Biomarkers Prev. 2009;18(1):275-81.
108. Lin JH, Zhang SM, Rexrode KM, Manson JE, Chan AT, Wu K, et al. Association between sex hormones and colorectal cancer risk in men and women. Clin Gastroenterol Hepatol. 2013;11(4):419-24.e1.
109. Murphy N, Strickler HD, Stanczyk FZ, Xue X, Wassertheil-Smoller S, Rohan TE, et al. A Prospective Evaluation of Endogenous Sex Hormone Levels and Colorectal Cancer Risk in Postmenopausal Women. J Natl Cancer Inst. 2015;107(10).

110. Henderson BE, Feigelson HS. Hormonal carcinogenesis. Carcinogenesis. 2000;21(3):427-33.
111. Key T, Appleby P, Barnes I, Reeves G, Endogenous H, Breast Cancer Collaborative G. Endogenous sex hormones and breast cancer in postmenopausal women: reanalysis of nine prospective studies. J Natl Cancer Inst. 2002;94(8):606-16.
112. Kaaks R, Lukanova A, Kurzer MS. Obesity, endogenous hormones, and endometrial cancer risk: a synthetic review. Cancer Epidemiol Biomarkers Prev. 2002;11(12):1531-43.
113. Allen NE, Key TJ, Dossus L, Rinaldi S, Cust A, Lukanova A, et al. Endogenous sex hormones and endometrial cancer risk in women in the European Prospective Investigation into Cancer and Nutrition (EPIC). Endocr Relat Cancer. 2008;15(2):485-97.
114. Brinton LA, Trabert B, Anderson GL, Falk RT, Felix AS, Fuhrman BJ, et al. Serum Estrogens and Estrogen Metabolites and Endometrial Cancer Risk among Postmenopausal Women. Cancer Epidemiol Biomarkers Prev. 2016;25(7):1081-9.
115. Lukanova A, Lundin E, Micheli A, Arslan A, Ferrari P, Rinaldi S, et al. Circulating levels of sex steroid hormones and risk of endometrial cancer in postmenopausal women. Int J Cancer. 2004;108(3):425-32.
116. Zeleniuch-Jacquotte A, Akhmedkhanov A, Kato I, Koenig KL, Shore RE, Kim MY, et al. Postmenopausal endogenous oestrogens and risk of endometrial cancer: results of a prospective study. Br J Cancer. 2001;84(7):975-81.
117. Aleksandrova K, Drogan D, Boeing H, Jenab M, Bas Bueno-de-Mesquita H, Jansen E, et al. Adiposity, mediating biomarkers and risk of colon cancer in the European prospective investigation into cancer and nutrition study. Int J Cancer. 2014;134(3):612-21.
118. Dossus L, Lukanova A, Rinaldi S, Allen N, Cust AE, Becker S, et al. Hormonal, metabolic, and inflammatory profiles and endometrial cancer risk within the EPIC cohort--a factor analysis. Am J Epidemiol. 2013;177(8):787-99.
119. Gunter MJ, Hoover DR, Yu H, Wassertheil-Smoller S, Rohan TE, Manson JE, et al. Insulin, insulin-like growth factor-I, and risk of breast cancer in postmenopausal women. J Natl Cancer Inst. 2009;101(1):48-60.
120. Hvidtfeldt UA, Gunter MJ, Lange T, Chlebowski RT, Lane D, Farhat GN, et al. Quantifying mediating effects of endogenous estrogen and insulin in the relation between obesity, alcohol consumption, and breast cancer. Cancer Epidemiol Biomarkers Prev. 2012;21(7):1203-12.
121. Rinaldi S, Key TJ, Peeters PH, Lahmann PH, Lukanova A, Dossus L, 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(11):2832-9.
122. Baron RM, Kenny DA. The moderator-mediator variable distinction in social psychological research: conceptual, strategic, and statistical considerations. J Pers Soc Psychol. 1986;51(6):1173-82.
123. Greenland S, Robins JM, Pearl J. Confounding and collapsibility in causal inference. Stat Sci. 1999;14(1):29-46.
124. Pearl J. Direct and Indirect effects. Proceedings of the seventeenth conference uncertain Artificial Intelligence. 2001:411-20.
125. Robins JM, Greenland S. Identifiability and exchangeability for direct and indirect effects. Epidemiology. 1992;3(2):143-55.
126. Rudolph KE, Goin DE, Paksarian D, Crowder R, Merikangas KR, Stuart EA. Causal Mediation Analysis With Observational Data: Considerations and Illustration Examining Mechanisms Linking Neighborhood Poverty to Adolescent Substance Use. Am J Epidemiol. 2019;188(3):598-608.

127. Vanderweele TJ, Vansteelandt S. Odds ratios for mediation analysis for a dichotomous outcome. Am J Epidemiol. 2010;172(12):1339-48.
128. VanderWeele TJ, Tchetgen Tchetgen EJ. Mediation Analysis With Matched Case-Control Study Designs. Am J Epidemiol. 2016;183(9):869-70.
129. VanderWeele TJ, Vansteelandt S. Mediation Analysis with Multiple Mediators. Epidemiol Methods. 2014;2(1):95-115.
130. Vansteelandt S, Daniel RM. Interventional Effects for Mediation Analysis with Multiple Mediators. Epidemiology. 2017;28(2):258-65.
131. Moreno-Betancur M, Carlin JB. Understanding Interventional Effects: A More Natural Approach to Mediation Analysis? Epidemiology. 2018;29(5):614-7.
132. Vanderweele TJ, Vansteelandt S, Robins JM. Effect decomposition in the presence of an exposure-induced mediator-outcome confounder. Epidemiology. 2014;25(2):300-6.
133. Petimar J, Tabung FK, Valeri L, Rosner B, Chan AT, Smith-Warner SA, et al. Mediation of associations between adiposity and colorectal cancer risk by inflammatory and metabolic biomarkers. Int J Cancer. 2019;144(12):2945-53.
134. Fritz J, Bjorge T, Nagel G, Manjer J, Engeland A, Haggstrom C, et al. The triglyceride-glucose index as a measure of insulin resistance and risk of obesity-related cancers. Int J Epidemiol. 2019.
135. Schairer C, Fuhrman BJ, Boyd-Morin J, Genkinger JM, Gail MH, Hoover RN, et al. Quantifying the Role of Circulating Unconjugated Estradiol in Mediating the Body Mass Index-Breast Cancer Association. Cancer Epidemiol Biomarkers Prev. 2016;25(1):105-13.
136. 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 Techn. 2012;1(3):123-6.
137. Milne RL, Fletcher AS, MacInnis RJ, Hodge AM, Hopkins AH, Bassett JK, et al. Cohort Profile: The Melbourne Collaborative Cohort Study (Health 2020). Int J Epidemiol. 2017;46(6):1757-i.
138. Baglietto L, Severi G, English DR, Krishnan K, Hopper JL, McLean C, et al. Circulating steroid hormone levels and risk of breast cancer for postmenopausal women. Cancer Epidemiol Biomarkers Prev. 2010;19(2):492-502.
139. Riboli E, Hunt KJ, Slimani N, Ferrari P, Norat T, Fahey M, et al. European Prospective Investigation into Cancer and Nutrition (EPIC): study populations and data collection. Public Health Nutr. 2002;5(6B):1113-24.
140. Cust AE, Kaaks R, Friedenreich C, Bonnet F, Laville M, Lukanova A, et al. Plasma adiponectin levels and endometrial cancer risk in pre- and postmenopausal women. J Clin Endocrinol Metab. 2007;92(1):255-63.
141. The Women's Health Initiative Study Group. Design of the Women's Health Initiative clinical trial and observational study. The Women's Health Initiative Study Group. Control Clin Trials. 1998;19(1):61-109.
142. UK Biobank. UK Biobank Biomarker Project Companion Document to Accompany Serum Biomarker Data: Version 1.0. http://biobank.ndph.ox.ac.uk/showcase/showcase/docs/serum_biochemistry.pdf; 2019.
143. UK Biobank. UK Biobank Biomarker Enhancement Project Companion Document to Accompany HbA1c Biomarker Data: Version 1.0. http://biobank.ndph.ox.ac.uk/showcase/showcase/docs/serum_hb1ac.pdf; 2018.
144. Lin SH, Young JG, Logan R, VanderWeele TJ. Mediation analysis for a survival outcome with time-varying exposures, mediators, and confounders. Stat Med. 2017;36(26):4153-66.
145. Samuelsen SO. A pseudolikelihood approach to analysis of nested case-control studies. Biometrika. 1997;84(2):379-94.
146. Schomaker M, Heumann H. Bootstrap inference when using multiple imputation. Statistics in Medicine. 2018;37(14):2252-66.

147. Murphy N, Jenab M, Gunter MJ. Adiposity and gastrointestinal cancers: epidemiology, mechanisms and future directions. Nat Rev Gastroenterol Hepatol. 2018;15(11):659-70.
148. Pelusi C, Pasquali R. The Significance of Low Testosterone Levels in Obese Men. Curr Obes Rep. 2012;1 (4):181–90.
149. Iyengar NM, Gucalp A, Dannenberg AJ, Hudis CA. Obesity and Cancer Mechanisms: Tumor Microenvironment and Inflammation. J Clin Oncol. 2016;34(35):4270-6.
150. Gambineri A, Pelusi C. Sex hormones, obesity and type 2 diabetes: is there a link? Endocr Connect. 2019;8(1):R1-R9.
151. UK Biobank. Body Composition Measurement: Version 1.0. http://biobank.ndph.ox.ac.uk/showcase/showcase/docs/body_composition.pdf; 2011.
152. UK Biobank. Biomarker assay quality procedures: approaches used to minimise systematic and random errors (and the wider epidemiological implications): Version 1.2. http://biobank.ndph.ox.ac.uk/showcase/showcase/docs/biomarker_issues.pdf; 2019.
153. Hosgood HD, Gunter MJ, Murphy N, Rohan TE, Strickler HD. The Relation of Obesity-Related Hormonal and Cytokine Levels With Multiple Myeloma and Non-Hodgkin Lymphoma. Front Oncol. 2018;8:103.
154. de Ronde W, van der Schouw YT, Pols HA, Gooren LJ, Muller M, Grobbee DE, et al. Calculation of bioavailable and free testosterone in men: a comparison of 5 published algorithms. Clin Chem. 2006;52(9):1777-84.
155. VanderWeele TJ. Mediation Analysis: A Practitioner's Guide. Annu Rev Public Health. 2016;37:17-32.
156. Cummings P. Methods for estimating adjusted risk ratios. Stata J. 2009;9(2):175-96.
157. Tilling K, Williamson EJ, Spratt M, Sterne JA, Carpenter JR. Appropriate inclusion of interactions was needed to avoid bias in multiple imputation. J Clin Epidemiol. 2016;80:107-15.
158. StataCorp. 2017. Stata Statistical Software: Release 15. College Station, TX: StataCorp LLC.
159. Ortega LS, Bradbury KE, Cross AJ, Morris JS, Gunter MJ, Murphy N. A Prospective Investigation of Body Size, Body Fat Composition and Colorectal Cancer Risk in the UK Biobank. Sci Rep. 2017;7(1):17807.
160. Pischon T, Lahmann PH, Boeing H, Friedenreich C, Norat T, Tjonneland A, et al. Body size and risk of colon and rectal cancer in the European Prospective Investigation Into Cancer and Nutrition (EPIC). J Natl Cancer Inst. 2006;98(13):920-31.
161. Murphy N, Ward HA, Jenab M, Rothwell JA, Boutron-Ruault MC, Carbonnel F, et al. Heterogeneity of Colorectal Cancer Risk Factors by Anatomical Subsite in 10 European Countries: A Multinational Cohort Study. Clin Gastroenterol Hepatol. 2019;17(7):1323-31 e6.
162. Xu J, Ye Y, Wu H, Duerksen-Hughes P, Zhang H, Li P, et al. Association between markers of glucose metabolism and risk of colorectal cancer. BMJ Open. 2016;6(6):e011430.
163. Fortner RT, Katzke V, Kuhn T, Kaaks R. Obesity and Breast Cancer. Recent Results Cancer Res. 2016;208:43-65.
164. Baglietto L, English DR, Hopper JL, MacInnis RJ, Morris HA, Tilley WD, et al. Circulating steroid hormone concentrations in postmenopausal women in relation to body size and composition. Breast cancer research and treatment. 2009;115(1):171-9.
165. Yager JD, Davidson NE. Estrogen carcinogenesis in breast cancer. N Engl J Med. 2006;354(3):270-82.
166. Kaaks R, Rinaldi S, Key TJ, Berrino F, Peeters PH, Biessy C, et al. Postmenopausal serum androgens, oestrogens and breast cancer risk: the European prospective investigation into cancer and nutrition. Endocr Relat Cancer. 2005;12(4):1071-82.

167. 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(16):1218-26.
168. Baglietto L, English DR, Hopper JL, Morris HA, Tilley WD, Giles GG. Circulating insulin-like growth factor-I and binding protein-3 and the risk of breast cancer. Cancer epidemiology, biomarkers & prevention : a publication of the American Association for Cancer Research, cosponsored by the American Society of Preventive Oncology. 2007;16(4):763-8.
169. Simpson JA, English DR, Macinnis RJ, Gertig DM, Hopper JL, Giles GG. A comparison of different methods for including 'age at menopause' in analyses of the association between hormone replacement therapy use and breast cancer. The journal of family planning and reproductive health care / Faculty of Family Planning & Reproductive Health Care, Royal College of Obstetricians & Gynaecologists. 2007;33(1):11-6.
170. Prentice RL. A Case-Cohort Design for Epidemiologic Cohort Studies and Disease Prevention Trials. Biometrika. 1986;73(1):1-11.
171. Xue X, Xie X, Gunter M, Rohan TE, Wassertheil-Smoller S, Ho GY, et al. Testing the proportional hazards assumption in case-cohort analysis. BMC Med Res Methodol. 2013;13:88.
172. Colditz, G., Baer, H., Rulla M.T. Breast Cancer. In (Ed.), Cancer Epidemiology and Prevention. : Oxford University Press,. Retrieved 4 Apr. 2019, from http://www.oxfordscholarship.com/view/10.1093/acprof:oso/9780195149616.001.0001/acprof-9780195149616-chapter-51. 2006.
173. Interventions IWGotEoC-P. Absence of excess body fatness, IARC Handbooks of Cancer Prevention ; Volume 16. 2018.
174. Breslow NE, Day NE. Statistical methods in cancer research. Volume I - The analysis of case-control studies. IARC Sci Publ. 1980(32):5-338.
175. StataCorp LP. Stata multiple-imputation reference manual : release 12. College Station, Tex.: Stata Press; 2011. iii, 365 p. p.
176. StataCorp. 2015. Stata Statistical Software: Release 14. College Station, TX: StataCorp LP.
177. Banack HR, Stokes A, Fox MP, Hovey KM, Cespedes Feliciano EM, LeBlanc ES, et al. Stratified Probabilistic Bias Analysis for Body Mass Index-related Exposure Misclassification in Postmenopausal Women. Epidemiology. 2018;29(5):604-13.
178. Falk RT, Gentzschein E, Stanczyk FZ, Garcia-Closas M, Figueroa JD, Ioffe OB, et al. Sex steroid hormone levels in breast adipose tissue and serum in postmenopausal women. Breast Cancer Res Treat. 2012;131(1):287-94.
179. 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. 2010;11(6):530-42.
180. Ferrannini E, Mari A. beta-Cell function in type 2 diabetes. Metabolism. 2014;63(10):1217-27.
181. Fader AN, Arriba LN, Frasure HE, von Gruenigen VE. Endometrial cancer and obesity: epidemiology, biomarkers, prevention and survivorship. Gynecol Oncol. 2009;114(1):121-7.
182. Shaw E, Farris M, McNeil J, Friedenreich C. Obesity and Endometrial Cancer. Recent Results Cancer Res. 2016;208:107-36.
183. Renehan AG, Zwahlen M, Egger M. Adiposity and cancer risk: new mechanistic insights from epidemiology. Nat Rev Cancer. 2015;15(8):484-98.
184. Calle EE, Kaaks R. Overweight, obesity and cancer: epidemiological evidence and proposed mechanisms. Nat Rev Cancer. 2004;4(8):579-91.
185. Merritt MA, Gunter MJ. Epidemiologic Evidence for the Obesity-Endometrial Cancer Relationship. Energ Balance Cancer. 2018;13:1-19.

186. Rinaldi S, Geay A, Dechaud H, Biessy C, Zeleniuch-Jacquotte A, Akhmedkhanov A, et al. Validity of free testosterone and free estradiol determinations in serum samples from postmenopausal women by theoretical calculations. Cancer Epidemiol Biomarkers Prev. 2002;11(10 Pt 1):1065-71.
187. Dossus L, Allen N, Kaaks R, Bakken K, Lund E, Tjonneland A, et al. Reproductive risk factors and endometrial cancer: the European Prospective Investigation into Cancer and Nutrition. Int J Cancer. 2010;127(2):442-51.
188. Cook, L., Weiss, N., Doherty, J., Chen, C. Endometrial Cancer. In (Ed.), Cancer Epidemiology and Prevention. : Oxford University Press,. Retrieved 7 Apr. 2019, from http://www.oxfordscholarship.com/view/10.1093/acprof:oso/9780195149616.001.0001/acprof-9780195149616-chapter-53. 2006.
189. Albert JM, Li Y, Sun J, Woyczynski WA, Nelson S. Continuous-time causal mediation analysis. Stat Med. 2019;38(22):4334-47.
190. Rodriguez AC, Blanchard Z, Maurer KA, Gertz J. Estrogen Signaling in Endometrial Cancer: a Key Oncogenic Pathway with Several Open Questions. Horm Cancer. 2019;10(2-3):51-63.
191. Chandran M, Phillips SA, Ciaraldi T, Henry RR. Adiponectin: more than just another fat cell hormone? Diabetes Care. 2003;26(8):2442-50.
192. Yoon YS, Kwon AR, Lee YK, Oh SW. Circulating adipokines and risk of obesity related cancers: A systematic review and meta-analysis. Obesity research & clinical practice. 2019;13(4):329-39.
193. Ho GY, Wang T, Zheng SL, Tinker L, Xu J, Rohan TE, et al. Circulating soluble cytokine receptors and colorectal cancer risk. Cancer Epidemiol Biomarkers Prev. 2014;23(1):179-88.
194. World Cancer Research Fund/American Institute for Cancer Research. Continuous Update Project Expect Report 2018. Diet, nutrition and physicial activity: Enegery balance and body fatness. Available at http://dietandcancerreport.org/.
195. Kaplan RC, Ho GY, Xue X, Rajpathak S, Cushman M, Rohan TE, et al. Within-individual stability of obesity-related biomarkers among women. Cancer Epidemiol Biomarkers Prev. 2007;16(6):1291-3.
196. Gunter MJ, Cross AJ, Huang WY, Stanczyk FZ, Purdue M, Xue X, et al. A prospective evaluation of C-reactive protein levels and colorectal adenoma development. Cancer Epidemiol Biomarkers Prev. 2011;20(3):537-44.
197. Hankinson SE, Manson JE, Spiegelman D, Willett WC, Longcope C, Speizer FE. Reproducibility of plasma hormone levels in postmenopausal women over a 2-3-year period. Cancer Epidemiol Biomarkers Prev. 1995;4(6):649-54.
198. Dashti SG, Simpson JA, Karahalios A, Viallon V, Moreno-Betancur M, Gurrin LC, et al. Adiposity and estrogen receptor-positive, postmenopausal breast cancer risk: Quantification of the mediating effects of fasting insulin and free estradiol. Int J Cancer. 2019.
199. Kim H, Giovannucci EL. Sex differences in the association of obesity and colorectal cancer risk. Cancer Causes Control. 2017;28(1):1-4.
200. le Cessie S. Bias Formulas for Estimating Direct and Indirect Effects When Unmeasured Confounding Is Present. Epidemiology. 2016;27(1):125-32.
201. Smith LH, VanderWeele TJ. Mediational E-values: Approximate Sensitivity Analysis for Unmeasured Mediator-Outcome Confounding. Epidemiology. 2019;30(6):835-7.

Appendices

Appendix A Supplementary material related to publication presented in Chapter 5

Supplementary Table A-1 Baseline Characteristics of men and women eligible for the study by categories of body mass index

	Men			Women		
Body Mass Index , kg/m2	≥18.5-<25	≥25-<30	≥30	≥18.5-<25	≥25-<30	≥30
	N=47,274	N=91,180	N=41,482	N=52,316	N=53,621	N=30,641
Colorectal cancer affected - yes, n (%)	279 (1)	666 (1)	335 (1)	269 (1)	339 (1)	182 (1)
Age at recruitment, median (interquartile range)	57.0 (49.0-63.0)	58.0 (50.0-63.0)	57.0 (50.0-63.0)	60.0 (56.0-64.0)	61.0 (57.0-65.0)	61.0 (57.0-64.0)
Age at cancer diagnosis or the end of follow up, median (25th - 75th percentiles)	64.0 (56.0-69.7)	64.9 (56.9-70.3)	64.3 (56.9-69.7)	67.6 (63.4-71.1)	68.2 (64.4-71.8)	67.9 (64.0-71.5)
Follow-up time - years, median (25th - 75th percentiles)	7.1 (6.5-7.7)	7.1 (6.4-7.7)	7.1 (6.4-7.8)	7.1 (6.4-7.7)	7.1 (6.5-7.7)	7.1 (6.4-7.7)
Highest educational attainment, n (%)						
None of the below	5,402 (11)	13,476 (15)	7,942 (19)	8,468 (16)	11,848 (22)	7,915 (26)
CSEs/O-levels/GCSEs or equivalent	11,778 (25)	21,085 (23)	9,253 (22)	14,386 (27)	14,244 (27)	7,672 (25)
NVQ/HND/HNC/A-levels/AS-levels or equivalent	8,810 (19)	19,013 (21)	9,199 (22)	6,135 (12)	6,877 (13)	4,393 (14)
Other professional quals. (eg nurse/teach)	13,722 (29)	25,754 (28)	10,912 (26)	16,839 (32)	15,632 (29)	8,314 (27)
College/uni degree	7,562 (16)	11,852 (13)	4,176 (10)	6,488 (12)	5,020 (9)	2,347 (8)
Townsend deprivation index at recruitment - quintiles, n (%)						
1	10,219 (22)	20,354 (22)	7,892 (19)	12,295 (24)	11,594 (22)	5,353 (17)
2	9,748 (21)	19,463 (21)	8,050 (19)	11,658 (22)	11,650 (22)	5,878 (19)
3	9,200 (19)	18,571 (20)	8,354 (20)	10,947 (21)	11,358 (21)	6,192 (20)
4	9,159 (19)	17,487 (19)	8,467 (20)	9,842 (19)	10,613 (20)	6,371 (21)
5	8,948 (19)	15,305 (17)	8,719 (21)	7,574 (14)	8,406 (16)	6,847 (22)
First degree relative history of colorectal cancer - yes, n (%)	4,981 (11)	10,491 (12)	4,933 (12)	6,153 (12)	6,500 (12)	3,726 (12)
Physical activity - MET hr/wk, n (%)						
<10 MET hr/wk	8,095 (17)	17,905 (20)	11,284 (27)	9,655 (18)	12,394 (23)	10,000 (33)
10-<20 MET hr/wk	7,743 (16)	15,507 (17)	7,129 (17)	9,224 (18)	9,965 (19)	5,773 (19)
20-<40 MET hr/wk	11,807 (25)	22,066 (24)	9,032 (22)	12,988 (25)	13,059 (24)	6,581 (21)
40-<60 MET hr/wk	6,887 (15)	12,105 (13)	4,566 (11)	7,402 (14)	6,763 (13)	3,192 (10)
60+ MET hr/wk	12,742 (27)	23,597 (26)	9,471 (23)	13,047 (25)	11,440 (21)	5,095 (17)
Alcohol intake, n (%)						
never	2,748 (6)	4,384 (5)	2,358 (6)	4,077 (8)	4,542 (8)	3,696 (12)
Current occasionally to 1-3 times/month	7,050 (15)	12,736 (14)	7,494 (18)	11,619 (22)	14,180 (26)	11,028 (36)
1-2 times/week	11,507 (24)	23,353 (26)	11,716 (28)	12,761 (24)	14,129 (26)	7,633 (25)
3-4 times/week	12,696 (27)	25,966 (28)	10,551 (25)	12,385 (24)	11,392 (21)	4,838 (16)
daily or almost daily	13,273 (28)	24,741 (27)	9,363 (23)	11,474 (22)	9,378 (17)	3,446 (11)
Smoking status, n (%)						
Never	26,764 (57)	45,942 (50)	18,644 (45)	31,789 (61)	31,062 (58)	17,433 (57)
Former	14,009 (30)	35,076 (38)	18,265 (44)	16,379 (31)	18,580 (35)	11,106 (36)
Current	6,501 (14)	10,162 (11)	4,573 (11)	4,148 (8)	3,979 (7)	2,102 (7)

continued

Supplementary Table A-1 continued Baseline Characteristics of men and women eligible for the study by categories of body mass index

	Men			Women		
Body Mass Index , kg/m2	≥18.5-<25	≥25-<30	≥30	≥18.5-<25	≥25-<30	≥30
	N=47,274	N=91,180	N=41,482	N=52,316	N=53,621	N=30,641
Total red and processed meat intake - times/week, median (25th - 75th percentiles)	3.5 (2.5-5.0)	4.0 (2.5-5.0)	4.5 (2.5-5.5)	2.5 (2.0-4.0)	2.5 (2.0-4.5)	3.0 (2.0-4.5)
Regular aspirin or ibuprofen use - yes, n (%)	9,833 (21)	24,035 (26)	13,676 (33)	10,100 (19)	12,990 (24)	8,947 (29)
Number of live births, median (25th - 75th percentiles)	-	-	-	2.0 (1.0-2.0)	2.0 (1.0-3.0)	2.0 (1.0-3.0)
Ever use of oral contraceptive pill - yes, n (%)	-	-	-	41,587 (79)	41,725 (78)	23,427 (76)
Ever use of hormone therapy - yes, n (%)	-	-	-	23,298 (45)	26,230 (49)	14,651 (48)
Adiposity measures						
Body Mass Index - kg/m2, median (25th - 75th percentiles)	23.5 (22.4-24.3)	27.3 (26.2-28.5)	32.2 (30.9-34.3)	23.0 (21.7-24.0)	27.1 (26.0-28.4)	32.9 (31.2-35.7)
Waist Circumference - cm, , n (%)						
≤80 F; ≤94 M	44,047 (93)	38,951 (43)	942 (2)	41,836 (80)	12,243 (23)	260 (1)
>80-≤88; F >94-≤102 M	3,134 (7)	39,700 (44)	8,394 (20)	9,267 (18)	23,690 (44)	2,863 (9)
>88 F; >102 M	93 (0)	12,529 (14)	32,146 (77)	1,213 (2)	17,688 (33)	27,518 (90)
Waist Circumference - cm, median (25th - 75th percentiles)	86.0 (82.0-90.0)	96.0 (92.0-100.0)	108.0 (103.0-114.0)	75.0 (71.0-79.0)	85.0 (81.0-90.0)	99.0 (93.0-105.0)
Waist-Hip ratio, n (%)						
≤0.77 F; ≤0.89 M	27,070 (57)	19,992 (22)	1,794 (4)	20,361 (39)	8,700 (16)	1,998 (7)
>0.77 - ≤0.86 F; >0.89 - ≤0.98 M	18,446 (39)	54,496 (60)	18,368 (44)	26,700 (51)	29,833 (56)	13,963 (46)
>0.86 F; >0.98 M	1,758 (4)	16,692 (18)	21,320 (51)	5,255 (10)	15,088 (28)	14,680 (48)
Waist-Hip Ratio, median (25th - 75th percentiles)	0.9 (0.8-0.9)	0.9 (0.9-1.0)	1.0 (0.9-1.0)	0.8 (0.7-0.8)	0.8 (0.8-0.9)	0.9 (0.8-0.9)
Biomarkers, median (25th - 75th percentiles)						
C-reactive protein - mg/L	0.8 (0.4-1.5)	1.2 (0.7-2.2)	2.0 (1.1-3.6)	0.8 (0.4-1.5)	1.5 (0.8-2.8)	3.0 (1.7-5.5)
C-reactive protein missing value - yes, n (%)	2,515 (5)	4,739 (5)	2,349 (6)	2,908 (6)	3,146 (6)	1,991 (6)
Glycated haemoglobin - nmol/mol	34.3 (32.1-36.5)	34.8 (32.5-37.2)	36.0 (33.4-38.7)	35.3 (33.2-37.4)	35.8 (33.6-38.0)	36.8 (34.4-39.3)
Glycated haemoglobin missing value - yes, n (%)	2,667 (6)	5,093 (6)	2,458 (6)	3,206 (6)	3,367 (6)	2,122 (7)
Sex Hormone Binding Globulin - nmol/L	44.0 (34.2-56.1)	36.3 (27.9-46.6)	31.9 (24.2-41.4)	67.2 (51.9-85.4)	51.8 (38.8-67.5)	40.2 (30.1-53.5)
Sex Hormone Binding Globulin missing value - yes, n (%)	6,336 (13)	11,765 (13)	5,328 (13)	7,680 (15)	8,043 (15)	4,602 (15)
Testosterone - pmol/L	13.1 (10.8-15.6)	11.7 (9.7-14.1)	10.4 (8.5-12.6)	0.9 (0.7-1.3)	1.0 (0.7-1.3)	1.0 (0.7-1.4)
Testosterone missing value - yes, n (%)	2,799 (6)	5,250 (6)	2,625 (6)	13,150 (25)	11,981 (22)	6,326 (21)
Missing for any of the biomarkers - yes, n (%)	8,702 (18)	16,256 (18)	7,469 (18)	18,663 (36)	17,794 (33)	9,660 (32)

CSE Certificate of Secondary Education; GCSE General Certificate of Secondary Education; NVQ National Vocational Education; HND Higher National Diploma; HNC Higher National Certificate; MET Metabolic Equivalent of Task; W women; M men

Supplementary Table A-2 Baseline Characteristics of men and women eligible for the study by categories of waist-hip ratio

	Men			Women		
Waist-Hip Ratio	≤0.89	>0.89 - ≤0.98	>0.98	≤0.77	>0.77 - ≤0.86	>0.86
	N=48,856	N=91,310	N=39,770	N=31,059	N=70,496	N=35,023
Colorectal cancer affected - yes, n (%)	234 (0)	680 (1)	366 (1)	144 (0)	412 (1)	234 (1)
Age at recruitment, median (interquartile range)	54.0 (47.0-61.0)	58.0 (50.0-63.0)	60.0 (53.0-64.0)	60.0 (56.0-64.0)	61.0 (57.0-64.0)	61.0 (58.0-65.0)
Age at cancer diagnosis or the end of follow up, median (25th - 75th percentiles)	61.5 (53.8-68.6)	64.8 (57.1-70.1)	66.8 (59.9-71.1)	67.0 (62.8-70.7)	68.0 (64.0-71.6)	68.5 (64.7-72.0)
Follow-up time - years, median (25th - 75th percentiles)	7.2 (6.5-7.8)	7.2 (6.5-7.8)	7.1 (6.4-7.7)	7.2 (6.5-7.7)	7.1 (6.5-7.7)	7.1 (6.4-7.7)
Highest educational attainment, n (%)						
None of the below	4,502 (9)	13,297 (15)	9,021 (23)	4,661 (15)	14,509 (21)	9,061 (26)
CSEs/O-levels/GCSEs or equivalent	12,168 (25)	21,372 (23)	8,576 (22)	8,605 (28)	18,775 (27)	8,922 (25)
NVQ/HND/HNC/A-levels/AS-levels or equivalent	9,679 (20)	19,191 (21)	8,152 (20)	3,816 (12)	8,971 (13)	4,618 (13)
Other professional quals. (eg nurse/teach)	14,680 (30)	25,748 (28)	9,960 (25)	10,226 (33)	21,095 (30)	9,464 (27)
College/uni degree	7,827 (16)	11,702 (13)	4,061 (10)	3,751 (12)	7,146 (10)	2,958 (8)
Townsend deprivation index at recruitment - quintiles, n (%)						
1	10,986 (22)	20,009 (22)	7,470 (19)	7,422 (24)	15,303 (22)	6,517 (19)
2	10,344 (21)	19,253 (21)	7,664 (19)	7,140 (23)	15,214 (22)	6,832 (20)
3	9,805 (20)	18,381 (20)	7,939 (20)	6,624 (21)	14,855 (21)	7,018 (20)
4	9,239 (19)	17,672 (19)	8,202 (21)	5,782 (19)	13,686 (19)	7,358 (21)
5	8,482 (17)	15,995 (18)	8,495 (21)	4,091 (13)	11,438 (16)	7,298 (21)
First degree relative history of colorectal cancer - yes, n (%)	5,057 (10)	10,425 (11)	4,923 (12)	3,570 (11)	8,460 (12)	4,349 (12)
Physical activity - MET hr/wk, n (%)						
<10 MET hr/wk	7,246 (15)	18,913 (21)	11,125 (28)	5,964 (19)	16,196 (23)	9,889 (28)
10-<20 MET hr/wk	7,527 (15)	15,729 (17)	7,123 (18)	5,600 (18)	12,889 (18)	6,473 (18)
20-<40 MET hr/wk	12,497 (26)	21,835 (24)	8,573 (22)	7,778 (25)	16,958 (24)	7,892 (23)
40-<60 MET hr/wk	7,520 (15)	11,779 (13)	4,259 (11)	4,347 (14)	8,973 (13)	4,037 (12)
60+ MET hr/wk	14,066 (29)	23,054 (25)	8,690 (22)	7,370 (24)	15,480 (22)	6,732 (19)
Alcohol intake, n (%)						
never	2,537 (5)	4,531 (5)	2,422 (6)	2,555 (8)	5,966 (8)	3,794 (11)
Current occasionally to 1-3 times/month	7,485 (15)	13,065 (14)	6,730 (17)	7,557 (24)	18,673 (26)	10,597 (30)
1-2 times/week	12,861 (26)	23,517 (26)	10,198 (26)	8,062 (26)	17,893 (25)	8,568 (24)
3-4 times/week	13,832 (28)	25,530 (28)	9,851 (25)	7,094 (23)	15,261 (22)	6,260 (18)
daily or almost daily	12,141 (25)	24,667 (27)	10,569 (27)	5,791 (19)	12,703 (18)	5,804 (17)
Smoking status, n (%)						
Never	29,345 (60)	45,477 (50)	16,528 (42)	20,152 (65)	41,571 (59)	18,561 (53)
Former	14,354 (29)	34,990 (38)	18,006 (45)	9,253 (30)	23,740 (34)	13,072 (37)
Current	5,157 (11)	10,843 (12)	5,236 (13)	1,654 (5)	5,185 (7)	3,390 (10)

continued

Supplementary Table A-2 continued Baseline Characteristics of men and women eligible for the study by categories of waist-hip ratio

	Men			Women		
Waist-Hip Ratio	≤0.89 N=48,856	>0.89 - ≤0.98 N=91,310	>0.98 N=39,770	≤0.77 N=31,059	>0.77 - ≤0.86 N=70,496	>0.86 N=35,023
Total red and processed meat intake - times/week, median (25th - 75th percentiles)	3.5 (2.5-5.0)	4.0 (2.5-5.0)	4.5 (2.5-5.5)	2.5 (2.0-4.0)	2.5 (2.0-4.5)	3.0 (2.0-4.5)
Regular aspirin or ibuprofen use - yes, n (%)	10,078 (21)	24,418 (27)	13,048 (33)	6,229 (20)	16,376 (23)	9,432 (27)
Number of live births, median (25th - 75th percentiles)	-	-	-	2.0 (1.0-2.0)	2.0 (1.0-3.0)	2.0 (1.0-3.0)
Ever use of oral contraceptive pill - yes, n (%)	-	-	-	24,839 (80)	55,206 (78)	26,694 (76)
Ever use of hormone therapy - yes, n (%)	-	-	-	13,478 (43)	33,734 (48)	16,967 (48)
Adiposity measures						
Waist-Hip Ratio, median (25th - 75th percentiles)	0.9 (0.8-0.9)	0.9 (0.9-1.0)	1.0 (1.0-1.0)	0.7 (0.7-0.8)	0.8 (0.8-0.8)	0.9 (0.9-0.9)
Body Mass Index - kg/m2, n (%)						
≥18.5-<25	27,070 (55)	18,446 (20)	1,758 (4)	20,361 (66)	26,700 (38)	5,255 (15)
≥25-<30	19,992 (41)	54,496 (60)	16,692 (42)	8,700 (28)	29,833 (42)	15,088 (43)
≥30	1,794 (4)	18,368 (20)	21,320 (54)	1,998 (6)	13,963 (20)	14,680 (42)
Body Mass Index - kg/m2, median (25th - 75th percentiles)	24.7 (23.0-26.5)	27.3 (25.4-29.4)	30.3 (28.0-33.1)	23.8 (21.9-26.0)	26.1 (23.8-29.1)	29.0 (26.3-32.5)
Waist Circumference - cm, n (%)						
≤80 F; ≤94 M	44,766 (92)	37,589 (41)	1,585 (4)	27,045 (87)	26,254 (37)	1,040 (3)
>80-≤88; F >94-≤102 M	3,681 (8)	37,444 (41)	10,103 (25)	3,140 (10)	25,958 (37)	6,722 (19)
>88 F; >102 M	409 (1)	16,277 (18)	28,082 (71)	874 (3)	18,284 (26)	27,261 (78)
Waist Circumference - cm, median (25th - 75th percentiles)	86.0 (82.0-90.0)	96.0 (92.0-101.0)	107.0 (102.0-113.0)	73.0 (69.0-77.0)	83.0 (78.0-89.0)	95.0 (89.0-102.0)
Biomarkers, median (25th - 75th percentiles)						
C-reactive protein - mg/L	0.8 (0.4-1.5)	1.3 (0.7-2.4)	1.9 (1.1-3.5)	0.8 (0.4-1.7)	1.4 (0.7-2.8)	2.2 (1.1-4.3)
C-reactive protein missing value - yes, n (%)	2,504 (5)	4,772 (5)	2,327 (6)	1,753 (6)	4,065 (6)	2,227 (6)
Glycated haemoglobin - nmol/mol	34.2 (32.0-36.3)	34.9 (32.6-37.3)	36.0 (33.5-38.7)	35.1 (33.0-37.1)	35.7 (33.6-38.0)	36.7 (34.4-39.2)
Glycated haemoglobin missing value - yes, n (%)	2,736 (6)	5,121 (6)	2,361 (6)	1,923 (6)	4,400 (6)	2,372 (7)
Sex Hormone Binding Globulin - nmol/L	41.5 (31.8-53.5)	36.2 (27.6-46.8)	34.0 (25.7-44.0)	69.0 (53.8-87.2)	54.4 (40.5-71.7)	42.9 (31.6-57.6)
Sex Hormone Binding Globulin missing value - yes, n (%)	6,408 (13)	11,834 (13)	5,187 (13)	4,625 (15)	10,368 (15)	5,332 (15)
Testosterone - pmol/L	12.9 (10.7-15.4)	11.7 (9.6-14.0)	10.6 (8.7-12.9)	1.0 (0.7-1.3)	1.0 (0.7-1.3)	1.0 (0.7-1.4)
Testosterone missing value - yes, n (%)	2,830 (6)	5,289 (6)	2,555 (6)	7,039 (23)	16,144 (23)	8,274 (24)
Missing for any of the biomarkers - yes, n (%)	8,828 (18)	16,375 (18)	7,224 (18)	10,438 (34)	23,686 (34)	11,993 (34)

CSE Certificate of Secondary Education; GCSE General Certificate of Secondary Education; NVQ National Vocational Education; HND Higher National Diploma; HNC Higher National Certificate; MET Metabolic Equivalent of Task; W women; M men

Supplementary Table A-3 Intraclass coefficients of biomarker levels measured at both recruitment and repeat assessment visit

	Men				Women			
Biomarker	No. with repeat measurements	ICC (95% CI)			No. with repeat measurements	ICC (95% CI)		
CRP	8318	0.56	(0.55 to	0.57)	8233	0.67	(0.66 to	0.68)
HbA1c	6299	0.84	(0.83 to	0.84)	6564	0.79	(0.79 to	0.80)
SHBG	6746	0.86	(0.86 to	0.87)	6384	0.81	(0.80 to	0.82)
Testosterone	8124	0.59	(0.58 to	0.61)	6064	0.63	(0.62 to	0.64)

Abbreviations CRP C-reactive protein; HbA1c glycated haemoglobin; SHBG sex hormone binding globulin

Supplementary Table A-4 Association between adiposity measures and biomarkers; complete case analyses excluding women in the middle category of each adiposity measure

		Ratio of geometric means (95% confidence interval)								
	Biomarker	Waist circumference >102 vs. ≤94 M >88 vs. ≤80 F			Body Mass Index ≥30 vs. <25 kg/m2			Waist-Hip Ratio >0.98 vs. ≤89 M >0.86 vs. ≤77 F		
Men		N=105,355			N=72,585			N=72,574		
	CRP	1.99	(1.96 to	2.01)	2.30	(2.26 to	2.33)	2.08	(2.05 to	2.11)
	HBA1C	1.04	(1.04 to	1.05)	1.05	(1.05 to	1.05)	1.05	(1.05 to	1.05)
	SHBG	0.79	(0.79 to	0.80)	0.72	(0.72 to	0.73)	0.77	(0.76 to	0.77)
	Testosterone	0.83	(0.83 to	0.83)	0.80	(0.80 to	0.81)	0.84	(0.83 to	0.84)
Women		N=66,718			N=54,634			N=43,651		
	CRP	2.67	(2.63 to	2.71)	3.25	(3.20 to	3.31)	2.21	(2.17 to	2.26)
	HBA1C	1.04	(1.04 to	1.04)	1.04	(1.04 to	1.04)	1.05	(1.05 to	1.05)
	SHBG	0.63	(0.63 to	0.64)	0.61	(0.61 to	0.61)	0.63	(0.62 to	0.63)
	Testosterone	1.08	(1.07 to	1.09)	1.13	(1.12 to	1.14)	1.04	(1.03 to	1.05)

Abbreviations: CRP C-reactive protein; HbA1c glycated haemoglobin; SHBG sex hormone binding globulin. Models were adjusted for age at recruitment, education, Townsend deprivation index at recruitment, first-degree relative history of colorectal cancer, physical activity, alcohol intake, smoking status, red and processed meat intake, and regular use of aspirin or ibuprofen. Models for women were additionally adjusted for the number of live births, ever use of oral contraceptive pill, and ever use of hormone therapy.

Supplementary Table A-5 Estimated odds ratios and 95% confidence intervals for the association between adiposity measures (exposure) or biomarkers (mediators) and colorectal cancer (outcome) in men and women complete case analysis

	Men				Women			
Waist Circumference, cm	**>102 vs <94**				**>88 vs <80**			
OR (95% CI)	1.38	(1.20 to	1.59)		1.28	(1.08 to	1.51)	
Body Mass Index, kg/m2	**≥30 vs <25**				**≥30 vs <25**			
OR (95% CI)	1.37	(1.16 to	1.62)		1.11	(0.91 to	1.36)	
Waist-Hip ratio	**>0.98 vs <0.89**				**>0.86 vs <0.77**			
OR (95% CI)	1.42	(1.19 to	1.70)		1.33	(1.07 to	1.65)	
Biomarker	**Per doubling concentration**			**P value for evidence against linearity**	**Per doubling concentration**			**P value for evidence against linearity**
CRP								
Model 1	1.06	(1.02 to	1.11)	0.01	1.03	(0.98 to	1.09)	0.63
HBA1C								
Model 1	0.86	(0.61 to	1.20)	0.13	1.09	(0.68 to	1.76)	0.75
Model 2 (additionally adjusted for CRP)	0.71	(0.50 to	1.00)	0.68	1.17	(0.72 to	1.89)	0.79
SHBG								
Model 1	0.90	(0.81 to	1.01)	0.19	1.01	(0.88 to	1.15)	0.00
Model 2 (additionally adjusted for CRP and HbA1c)	0.87	(0.77 to	0.98)	0.20	1.02	(0.88 to	1.17)	0.00
Testosterone								
Model 1	0.92	(0.81 to	1.05)	0.51	1.01	(0.90 to	1.14)	0.11
Model 2 (additionally adjusted for CRP and HbA1c)	0.90	(0.79 to	1.03)	0.47	1.00	(0.89 to	1.13)	0.08

Abbreviations: CRP C-reactive protein; HbA1c glycated haemoglobin; SHBG sex hormone binding globulin
Models were adjusted for age at recruitment, education, Townsend deprivation index at recruitment, first-degree relative history of colorectal cancer, physical activity, alcohol intake, smoking status, red and processed meat intake, and regular use of aspirin or ibuprofen. Models for women were additionally adjusted for the number of live births, ever use of oral contraceptive pill, and ever use of hormone therapy. Model 1 for biomarkers were additionally adjusted for waist circumference.

Supplementary Table A-6 Estimated natural direct and indirect effects using sequential mediation analysis for the association between adiposity measures and colorectal cancer for men and women, analyses exclude participants categorized as overweight and do not include any interaction terms between adiposity measures and biomarkers (exposure-mediator interactions) or biomarkers (mediator-mediator interactions) - complete case analysis

		Risk Ratio (95% CI)					
	Effect	Men			Women		
Waist Circumference >102 vs. ≤94 cm M >88 vs. ≤80 cm F	No. Colorectal Cancer affected/ No. unaffected	726/104,629			395/66,323		
	Total effect	1.47	(1.27 to	1.72)	1.12	(0.91 to	1.38)
	Natural indirect effect through all the mediators	1.11	(1.04 to	1.19)	1.03	(0.90 to	1.19)
	Natural indirect effect through CRP	1.07	(1.01 to	1.12)	1.04	(0.93 to	1.16)
	Natural indirect effect through HbA1C, excluding the influence of CRP	0.99	(0.97 to	1.01)	1.00	(0.97 to	1.03)
	Natural indirect effect through SHBG and testosterone, excluding the influence of CRP and HbA1c	1.06	(1.01 to	1.10)	1.00	(0.90 to	1.11)
	Natural direct effect not through any of the mediators	1.33	(1.12 to	1.56)	1.08	(0.83 to	1.41)
Body Mass Index ≥30 vs. <25 kg/m2	No. Colorectal Cancer affected/ No. unaffected	501/72,084			298/54,336		
	Total effect	1.49	(1.23 to	1.82)	1.05	(0.81 to	1.35)
	Natural indirect effect through all the mediators	1.12	(1.01 to	1.24)	-	-	-
	Natural indirect effect through CRP	1.07	(1.00 to	1.15)	-	-	-
	Natural indirect effect through HbA1C, excluding the influence of CRP	0.98	(0.95 to	1.01)	-	-	-
	Natural indirect effect through SHBG and testosterone, excluding the influence of CRP and HbA1c	1.07	(1.00 to	1.14)	-	-	-
	Natural direct effect not through any of the mediators	1.33	(1.08 to	1.65)	-	-	-
Waist-Hi Ratio >0.98 vs. <0.89 M >0.86 vs. ≤0.77 F	No. Colorectal Cancer affected/ No. unaffected	491/72,083			255/43,396		
	Total effect	1.37	(1.12 to	1.69)	1.24	(0.95 to	1.63)
	Natural indirect effect through all the mediators	1.13	(1.03 to	1.25)	1.05	(0.90 to	1.23)
	Natural indirect effect through CRP	1.07	(1.01 to	1.15)	1.04	(0.94 to	1.16)
	Natural indirect effect through HbA1C, excluding the influence of CRP	0.98	(0.95 to	1.01)	1.01	(0.97 to	1.06)
	Natural indirect effect through SHBG and testosterone, excluding the influence of CRP and HbA1c	1.08	(1.02 to	1.14)	1.00	(0.89 to	1.11)
	Natural direct effect not through any of the mediators	1.21	(0.97 to	1.52)	1.18	(0.86 to	1.62)

Abbreviations: CRP C-reactive protein; HbA1c glycated haemoglobin; SHBG sex hormone binding globulin
Models were adjusted for age at recruitment, education, Townsend deprivation index at recruitment, first-degree relative history of colorectal cancer, physical activity, alcohol intake, smoking status, red and processed meat intake, and regular use of aspirin or ibuprofen. Models for women were additionally adjusted for the number of live births, ever use of oral contraceptives, and ever use of hormone therapy.

Supplementary Table A-7 Estimated natural direct and indirect effects using sequential mediation analysis for the association between adiposity measures and colorectal cancer for men and women, analyses exclude participants categorized as overweight and include all interaction terms between adiposity measures and biomarkers (exposure-mediator interactions) and between biomarkers (mediator-mediator interactions)

		Men						Women					
		Risk Ratio (95% CI)						Risk Ratio (95% CI)					
	Effect	Complete case analysis			Multiple imputation			Complete case analysis			Multiple imputation		
Waist Circumference >102 vs. ≤94 cm M >88 vs. ≤80 cm F	No. Colorectal Cancer affected/ No. unaffected	430/104,629			516/126,602			294/66,323			448/98,655		
	Total effect	1.47	(1.27 to	1.72)	1.37	(1.19 to	1.58)	1.12	(0.91 to	1.38)	1.27	(1.07 to	1.50)
	Natural indirect effect through all the mediators	1.09	(0.98 to	1.21)	1.07	(0.97 to	1.18)	1.01	(0.79 to	1.29)	1.09	(0.89 to	1.33)
	Natural indirect effect through CRP	1.05	(0.98 to	1.14)	1.06	(0.98 to	1.13)	0.99	(0.84 to	1.16)	1.06	(0.94 to	1.20)
	Natural indirect effect through HbA1C, excluding the influence of CRP	0.99	(0.96 to	1.02)	0.99	(0.96 to	1.02)	1.06	(1.01 to	1.11)	1.05	(1.02 to	1.09)
	Natural indirect effect through SHBG and testosterone, excluding the influence of CRP and HbA1c	1.05	(0.98 to	1.12)	1.02	(0.95 to	1.09)	0.96	(0.80 to	1.16)	0.97	(0.83 to	1.14)
	Natural direct effect not through any of the mediators	1.35	(1.12 to	1.63)	1.28	(1.08 to	1.52)	1.11	(0.80 to	1.55)	1.16	(0.90 to	1.51)
Body Mass Index ≥30 vs. <25 kg/m2	No. Colorectal Cancer affected/ No. unaffected	294/72,084			359/87,265			222/54,336			329/81,307		
	Total effect	1.49	(1.23 to	1.82)	1.36	(1.15 to	1.60)	1.05	(0.81 to	1.35)	1.12	(0.92 to	1.38)
	Natural indirect effect through all the mediators	1.16	(1.00 to	1.34)	1.11	(0.96 to	1.28)	-	-	-	-	-	-
	Natural indirect effect through CRP	1.07	(0.98 to	1.18)	1.07	(0.98 to	1.17)	-	-	-	-	-	-
	Natural indirect effect through HbA1C, excluding the influence of CRP	1.00	(0.96 to	1.04)	1.00	(0.96 to	1.04)	-	-	-	-	-	-
	Natural indirect effect through SHBG and testosterone, excluding the influence of CRP and HbA1c	1.08	(0.97 to	1.21)	1.03	(0.93 to	1.15)	-	-	-	-	-	-
	Natural direct effect not through any of the mediators	1.29	(1.01 to	1.65)	1.22	(0.99 to	1.52)	-	-	-	-	-	-

continued

Supplementary Table A-7 continued Estimated natural direct and indirect effects using sequential mediation analysis for the association between adiposity measures and colorectal cancer for men and women, analyses exclude participants categorized as overweight and include all interaction terms between adiposity measures and biomarkers (exposure-mediator interactions) and between biomarkers (mediator-mediator interactions)

		Men						Women					
		Risk Ratio (95% CI)						Risk Ratio (95% CI)					
	Effect	Complete case analysis			Multiple imputation			Complete case analysis			Multiple imputation		
	No. Colorectal Cancer affected/ No. unaffected	289/72,083			352/87,169			199/43,396			279/64,671		
	Total effect	1.37	(1.12 to	1.69)	1.31	(1.09 to	1.58)	1.24	(0.95 to	1.63)	1.33	(1.06 to	1.66)
	Natural indirect effect through all the mediators	1.06	(0.94 to	1.20)	1.05	(0.94 to	1.18)	1.01	(0.81 to	1.26)	1.12	(0.93 to	1.35)
	Natural indirect effect through CRP	1.09	(1.01 to	1.19)	1.09	(1.01 to	1.18)	1.01	(0.88 to	1.16)	1.06	(0.95 to	1.19)
	Natural indirect effect through HbA1C, excluding the influence of CRP	0.97	(0.93 to	1.01)	0.98	(0.95 to	1.02)	1.04	(0.97 to	1.12)	1.04	(0.99 to	1.10)
	Natural indirect effect through SHBG and testosterone, excluding the influence of CRP and HbA1c	1.00	(0.92 to	1.09)	0.98	(0.90 to	1.06)	0.96	(0.81 to	1.14)	1.01	(0.86 to	1.18)
Waist-Hip Ratio >0.98 vs. <0.89 M >0.86 vs. ≤0.77 cm F	Natural direct effect not through any of the mediators	1.29	(1.02 to	1.63)	1.25	(1.01 to	1.55)	1.23	(0.87 to	1.74)	1.18	(0.89 to	1.58)

Abbreviations: CRP C-reactive protein; HbA1c glycated haemoglobin; SHBG sex hormone binding globulin

Models were adjusted for age at recruitment, education, Townsend deprivation index at recruitment, first-degree relative history of colorectal cancer, physical activity, alcohol intake, smoking status, red and processed meat intake, and regular use of aspirin or ibuprofen. Models for women were additionally adjusted for the number of live births, ever use of oral contraceptives, and ever use of hormone therapy.

Supplementary Table A-8 Estimated natural direct and indirect effects using sequential mediation analysis for the association between adiposity measures and colorectal cancer for men and women, analyses exclude participants categorized as overweight, calculated free testosterone (instead of SHBG and testosterone) has been used as the biomarker on sex-steroid hormone pathway for this analysis

		Men						Women					
		Risk Ratio (95% CI)						Risk Ratio (95% CI)					
	Effect	Complete case analysis			Multiple imputation			Complete case analysis			Multiple imputation		
Waist Circumference >102 vs. ≤94 cm M >88 vs. ≤80 cm F	No. Colorectal Cancer affected/ No. unaffected	726/104,594			878/127,830			395/66,299			616/100,142		
	Total effect	1.47	(1.27 to	1.71)	1.37	(1.19 to	1.58)	1.12	(0.91 to	1.37)	1.27	(1.07 to	1.50)
	Natural indirect effect through all the mediators	1.06	(1.00 to	1.12)	1.05	(1.00 to	1.11)	1.04	(0.92 to	1.18)	1.07	(0.97 to	1.19)
	Natural indirect effect through CRP	1.07	(1.01 to	1.12)	1.06	(1.01 to	1.11)	1.04	(0.93 to	1.16)	1.08	(0.99 to	1.17)
	Natural indirect effect through HbA1C, excluding the influence of CRP	0.99	(0.97 to	1.01)	0.99	(0.97 to	1.01)	1.00	(0.97 to	1.02)	1.00	(0.98 to	1.02)
	Natural indirect effect through calculated free testosterone, excluding the influence of CRP and HbA1c	1.01	(0.99 to	1.02)	1.01	(0.99 to	1.02)	1.01	(0.95 to	1.07)	1.00	(0.94 to	1.05)
	Natural direct effect not through any of the mediators	1.39	(1.19 to	1.63)	1.30	(1.12 to	1.51)	1.08	(0.85 to	1.37)	1.18	(0.97 to	1.44)
Body Mass Index ≥30 vs. <25 kg/m2	No. Colorectal Cancer affected/ No. unaffected	501/72,052			614/88,142			298/54,316			451/82,506		
	Total effect	1.49	(1.22 to	1.82)	1.36	(1.15 to	1.60)	1.05	(0.81 to	1.35)	1.12	(0.92 to	1.38)
	Natural indirect effect through all the mediators	1.06	(0.98 to	1.14)	1.06	(0.98 to	1.14)	-	-	-	-	-	-
	Natural indirect effect through CRP	1.07	(1.00 to	1.15)	1.07	(1.00 to	1.14)	-	-	-	-	-	-
	Natural indirect effect through HbA1C, excluding the influence of CRP	0.98	(0.95 to	1.01)	0.99	(0.96 to	1.02)	-	-	-	-	-	-
	Natural indirect effect through calculated free testosterone, excluding the influence of CRP and HbA1c	1.01	(0.99 to	1.02)	1.01	(0.99 to	1.02)	-	-	-	-	-	-
	Natural direct effect not through any of the mediators	1.41	(1.15 to	1.74)	1.28	(1.07 to	1.54)	-	-	-	-	-	-

continued

Supplementary Table A-8 continued Estimated natural direct and indirect effects using sequential mediation analysis for the association between adiposity measures and colorectal cancer for men and women, analyses exclude participants categorized as overweight, calculated free testosterone (instead of SHBG and testosterone) has been used as the biomarker on sex-steroid hormone pathway for this analysis

		Men						Women					
		Risk Ratio (95% CI)						Risk Ratio (95% CI)					
	Effect	Complete case analysis			Multiple imputation			Complete case analysis			Multiple imputation		
	No. Colorectal Cancer affected/ No. unaffected	491/72,055			600/88,026			255/ 43,384			378/65,704		
	Total effect	1.37	(1.13 to	1.67)	1.31	(1.09 to	1.58)	1.24	(0.95 to	1.63)	1.33	(1.06 to	1.66)
	Natural indirect effect through all the mediators	1.06	(0.98 to	1.14)	1.07	(0.99 to	1.15)	1.05	(0.93 to	1.19)	1.09	(0.98 to	1.21)
	Natural indirect effect through CRP	1.07	(1.00 to	1.15)	1.07	(1.00 to	1.15)	1.04	(0.94 to	1.16)	1.09	(1.00 to	1.19)
	Natural indirect effect through HbA1C, excluding the influence of CRP	0.98	(0.95 to	1.00)	0.99	(0.96 to	1.02)	1.01	(0.97 to	1.06)	1.01	(0.97 to	1.05)
	Natural indirect effect through calculated free testosterone, excluding the influence of CRP and HbA1c	1.01	(1.00 to	1.01)	1.00	(1.00 to	1.01)	0.99	(0.93 to	1.06)	0.99	(0.94 to	1.05)
Waist-Hip Ratio >0.98 vs. <0.89 M >0.86 vs. ≤0.77 cm F	Natural direct effect not through any of the mediators	1.30	(1.05 to	1.61)	1.23	(1.01 to	1.50)	1.18	(0.88 to	1.59)	1.22	(0.95 to	1.57)

Abbreviations: CRP C-reactive protein; HbA1c glycated haemoglobin; SHBG sex hormone binding globulin
Models were adjusted for age at recruitment, education, Townsend deprivation index at recruitment, first degree relative history of colorectal cancer, physical activity, alcohol intake, smoking status, red and processed meat intake, and regular use of aspirin or ibuprofen. Models for women were additionally adjusted for number of live births, ever use of oral contraceptives, and ever use of hormone therapy.

Supplementary Table A-9 Estimated natural direct and indirect effects using sequential mediation analysis for the association between adiposity measures and colon cancer for men and women, analyses exclude participants categorized as overweight and do not include any interaction terms between adiposity measures and biomarkers (exposure-mediator interactions) or biomarkers (mediator-mediator interactions)

		Men						Women					
		Risk Ratio (95% CI)						Risk Ratio (95% CI)					
	Effect	Complete case analysis			Multiple imputation			Complete case analysis			Multiple imputation		
Waist Circumference >102 vs. ≤94 cm M >88 vs. ≤80 cm F	No. Colon Cancer affected/ No. unaffected	726/104,629			872/126,602			395/66,323			610/98,655		
	Total effect	1.71	(1.40 to	2.08)	1.53	(1.29 to	1.82)	1.16	(0.91 to	1.48)	1.31	(1.08 to	1.60)
	Natural indirect effect through all the mediators	1.16	(1.06 to	1.26)	1.12	(1.03 to	1.22)	1.04	(0.89 to	1.22)	1.06	(0.93 to	1.21)
	Natural indirect effect through CRP	1.06	(0.99 to	1.13)	1.05	(0.99 to	1.12)	1.02	(0.90 to	1.16)	1.07	(0.97 to	1.17)
	Natural indirect effect through HbA1C, excluding the influence of CRP	1.00	(0.97 to	1.03)	1.00	(0.98 to	1.03)	0.99	(0.96 to	1.02)	0.99	(0.96 to	1.02)
	Natural indirect effect through SHBG and testosterone, excluding the influence of CRP and HbA1c	1.09	(1.04 to	1.15)	1.06	(1.01 to	1.12)	1.03	(0.92 to	1.15)	1.00	(0.91 to	1.10)
	Natural direct effect not through any of the mediators	1.47	(1.19 to	1.83)	1.37	(1.12 to	1.66)	1.11	(0.83 to	1.50)	1.24	(0.98 to	1.57)
Body Mass Index ≥30 vs. <25 kg/m2	No. Colon Cancer affected/ No. unaffected	501/72,084			607/87,265			298/54,336			449/81,307		
	Total effect	1.82	(1.40 to	2.37)	1.60	(1.28 to	2.00)	1.15	(0.86 to	1.53)	1.23	(0.97 to	1.56)
	Natural indirect effect through all the mediators	1.18	(1.04 to	1.35)	1.16	(1.03 to	1.31)	-	-	-	-	-	-
	Natural indirect effect through CRP	1.08	(0.99 to	1.18)	1.07	(0.99 to	1.16)	-	-	-	-	-	-
	Natural indirect effect through HbA1C, excluding the influence of CRP	1.00	(0.96 to	1.04)	1.00	(0.97 to	1.04)	-	-	-	-	-	-
	Natural indirect effect through SHBG and testosterone, excluding the influence of CRP and HbA1c	1.10	(1.01 to	1.20)	1.08	(0.99 to	1.17)	-	-	-	-	-	-
	Natural direct effect not through any of the mediators	1.54	(1.15 to	2.07)	1.37	(1.06 to	1.78)	-	-	-	-	-	-

continued

Supplementary Table A-9 continued Estimated natural direct and indirect effects using sequential mediation analysis for the association between adiposity measures and colon cancer for men and women, analyses exclude participants categorized as overweight and do not include any interaction terms between adiposity measures and biomarkers (exposure-mediator interactions) or biomarkers (mediator-mediator interactions)

		Men						Women					
		Risk Ratio (95% CI)						Risk Ratio (95% CI)					
	Effect	Complete case analysis			Multiple imputation			Complete case analysis			Multiple imputation		
Waist-Hip Ratio >0.98 vs. <0.89 M >0.86 vs. ≤0.77 cm F	No. Colon Cancer affected/ No. unaffected	491/72,083			597/87,169			255/43,396			374/64,671		
	Total effect	1.54	(1.16 to	2.03)	1.46	(1.14 to	1.87)	1.35	(0.99 to	1.83)	1.39	(1.07 to	1.81)
	Natural indirect effect through all the mediators	1.11	(0.99 to	1.25)	1.09	(0.98 to	1.22)	1.06	(0.89 to	1.26)	1.08	(0.93 to	1.25)
	Natural indirect effect through CRP	1.06	(0.97 to	1.16)	1.05	(0.97 to	1.14)	1.03	(0.91 to	1.16)	1.07	(0.97 to	1.18)
	Natural indirect effect through HbA1C, excluding the influence of CRP	0.98	(0.95 to	1.02)	1.00	(0.96 to	1.03)	1.01	(0.96 to	1.06)	1.01	(0.96 to	1.05)
	Natural indirect effect through SHBG and testosterone, excluding the influence of CRP and HbA1c	1.07	(1.00 to	1.14)	1.04	(0.98 to	1.12)	1.02	(0.90 to	1.16)	1.00	(0.90 to	1.12)
	Natural direct effect not through any of the mediators	1.38	(1.01 to	1.89)	1.33	(1.01 to	1.76)	1.27	(0.89 to	1.82)	1.29	(0.96 to	1.74)

Abbreviations: CRP C-reactive protein; HbA1c glycated haemoglobin; SHBG sex hormone binding globulin

Models were adjusted for age at recruitment, education, Townsend deprivation index at recruitment, first-degree relative history of colorectal cancer, physical activity, alcohol intake, smoking status, red and processed meat intake, and regular use of aspirin or ibuprofen. Models for women were additionally adjusted for the number of live births, ever use of oral contraceptives, and ever use of hormone therapy.

Supplementary Table A-10 Estimated natural direct and indirect effects using sequential mediation analysis for the association between adiposity measures and colorectal cancer for men and women, analyses exclude participants categorized as overweight and do not include any interaction terms between adiposity measures and biomarkers (exposure-mediator interactions) or biomarkers (mediator-mediator interactions), analysis limited to colorectal cancer cases diagnosed after two years of follow-up and non-cases with more than two years of follow-up

		Men						Women					
		Risk Ratio (95% CI)						Risk Ratio (95% CI)					
	Effect	Complete case analysis			Multiple imputation			Complete case analysis			Multiple imputation		
Waist Circumference >102 vs. ≤94 cm M >88 vs. ≤80 cm F	No. Colon Cancer affected/ No. unaffected	609/104,503			730/127,665			341/66,254			519/100,042		
	Total effect	1.54	(1.30 to	1.82)	1.44	(1.24 to	1.66)	1.12	(0.90 to	1.41)	1.29	(1.07 to	1.55)
	Natural indirect effect through all the mediators	1.12	(1.04 to	1.21)	1.08	(1.01 to	1.16)	1.04	(0.89 to	1.22)	1.06	(0.93 to	1.21)
	Natural indirect effect through CRP	1.06	(1.00 to	1.12)	1.05	(0.99 to	1.10)	1.03	(0.92 to	1.16)	1.06	(0.96 to	1.16)
	Natural indirect effect through HbA1C, excluding the influence of CRP	0.99	(0.97 to	1.02)	0.99	(0.97 to	1.02)	1.00	(0.97 to	1.03)	1.01	(0.98 to	1.03)
	Natural indirect effect through SHBG and testosterone, excluding the influence of CRP and HbA1c	1.07	(1.02 to	1.12)	1.04	(1.00 to	1.09)	1.00	(0.91 to	1.11)	1.00	(0.91 to	1.09)
	Natural direct effect not through any of the mediators	1.37	(1.14 to	1.65)	1.32	(1.13 to	1.55)	1.08	(0.82 to	1.42)	1.21	(0.97 to	1.52)
Body Mass Index ≥30 vs. <25 kg/m2	No. Colon Cancer affected/ No. unaffected	425/71,982			523/88,010			265/54,283			395/82,423		
	Total effect	1.54	(1.25 to	1.89)	1.39	(1.15 to	1.69)	1.11	(0.86 to	1.45)	1.16	(0.93 to	1.46)
	Natural indirect effect through all the mediators	1.13	(1.01 to	1.26)	1.12	(1.00 to	1.24)	-	-	-	-	-	-
	Natural indirect effect through CRP	1.08	(1.00 to	1.16)	1.07	(0.99 to	1.14)	-	-	-	-	-	-
	Natural indirect effect through HbA1C, excluding the influence of CRP	0.99	(0.96 to	1.02)	0.99	(0.96 to	1.03)	-	-	-	-	-	-
	Natural indirect effect through SHBG and testosterone, excluding the influence of CRP and HbA1c	1.06	(0.99 to	1.14)	1.05	(0.98 to	1.13)	-	-	-	-	-	-
	Natural direct effect not through any of the mediators	1.36	(1.09 to	1.70)	1.25	(1.00 to	1.55)	-	-	-	-	-	-

continued

Supplementary Table A-10 continued Estimated natural direct and indirect effects using sequential mediation analysis for the association between adiposity measures and colorectal cancer for men and women, analyses exclude participants categorized as overweight and do not include any interaction terms between adiposity measures and biomarkers (exposure-mediator interactions) or biomarkers (mediator-mediator interactions), analysis limited to colorectal cancer cases diagnosed after two years of follow-up and non-cases with more than two years of follow-up

		Men						Women					
		Risk Ratio (95% CI)						Risk Ratio (95% CI)					
	Effect	Complete case analysis			Multiple imputation			Complete case analysis			Multiple imputation		
Waist-Hip Ratio >0.98 vs. <0.89 M >0.86 vs. ≤0.77 cm F	No. Colon Cancer affected/ No. unaffected	421/71,990			514/87,905			218/43,355			319/65,634		
	Total effect	1.43	(1.14 to	1.78)	1.41	(1.16 to	1.72)	1.16	(0.88 to	1.55)	1.30	(1.03 to	1.63)
	Natural indirect effect through all the mediators	1.14	(1.03 to	1.26)	1.12	(1.02 to	1.23)	1.03	(0.87 to	1.22)	1.06	(0.92 to	1.22)
	Natural indirect effect through CRP	1.08	(1.01 to	1.16)	1.07	(1.00 to	1.15)	1.04	(0.93 to	1.17)	1.06	(0.97 to	1.17)
	Natural indirect effect through HbA1C, excluding the influence of CRP	0.98	(0.96 to	1.01)	0.99	(0.96 to	1.02)	1.02	(0.98 to	1.07)	1.02	(0.98 to	1.06)
	Natural indirect effect through SHBG and testosterone, excluding the influence of CRP and HbA1c	1.07	(1.01 to	1.14)	1.05	(1.00 to	1.11)	0.97	(0.85 to	1.10)	0.98	(0.88 to	1.09)
	Natural direct effect not through any of the mediators	1.25	(0.97 to	1.61)	1.26	(1.01 to	1.57)	1.13	(0.80 to	1.58)	1.22	(0.94 to	1.59)

Abbreviations: CRP C-reactive protein; HbA1c glycated haemoglobin; SHBG sex hormone binding globulin

Models were adjusted for age at recruitment, education, Townsend deprivation index at recruitment, first-degree relative history of colorectal cancer, physical activity, alcohol intake, smoking status, red and processed meat intake, and regular use of aspirin or ibuprofen. Models for women were additionally adjusted for the number of live births, ever use of oral contraceptives, and ever use of hormone therapy.

Supplementary Table A-11 Distribution of baseline characteristics by missing status for any biomarker data for men and women

	Men		Women	
	No Missing for biomarker data N=147,509	Any missing for biomarker data N=32,427	No Missing for biomarker data N=90,461	Any missing for biomarker data N=46,117
Colorectal cancer affected - yes, n (%)	1,059 (1)	221 (1)	512 (1)	278 (1)
Age at recruitment, median (25th - 75th percentiles)	57.0 (50.0-63.0)	57.0 (50.0-63.0)	61.0 (57.0-64.0)	61.0 (57.0-65.0)
Age at cancer diagnosis or the end of follow up, median (25th - 75th percentiles)	64.5 (56.7-70.0)	64.5 (56.6-69.9)	67.8 (63.7-71.4)	68.1 (64.2-71.7)
Follow-up time - years, median (25th - 75th percentiles)	7.2 (6.5-7.8)	7.1 (6.4-7.7)	7.2 (6.5-7.7)	7.1 (6.4-7.7)
Highest educational attainment, n (%)				
None of the above	22,026 (15)	4,794 (15)	18,553 (21)	9,678 (21)
CSEs/O-levels/GCSEs or equivalent	34,583 (23)	7,533 (23)	24,199 (27)	12,103 (26)
NVQ/HND/HNC/A-levels/AS-levels or equivalent	30,418 (21)	6,604 (20)	11,568 (13)	5,837 (13)
Other professional quals. (eg nurse/teach)	41,266 (28)	9,122 (28)	26,918 (30)	13,867 (30)
College/uni degree	19,216 (13)	4,374 (13)	9,223 (10)	4,632 (10)
Townsend deprivation index at recruitment - quintiles, n (%)				
1	31,671 (21)	6,794 (21)	19,518 (22)	9,724 (21)
2	30,668 (21)	6,593 (20)	19,415 (21)	9,771 (21)
3	29,763 (20)	6,362 (20)	18,856 (21)	9,641 (21)
4	28,731 (19)	6,382 (20)	17,805 (20)	9,021 (20)
5	26,676 (18)	6,296 (19)	14,867 (16)	7,960 (17)
First degree relative history of colorectal cancer - yes, n (%)	16,849 (11)	3,556 (11)	10,825 (12)	5,554 (12)
Physical activity - MET hr/wk, n (%)				
<10 MET hr/wk	30,497 (21)	6,787 (21)	21,354 (24)	10,695 (23)
10-<20 MET hr/wk	24,845 (17)	5,534 (17)	16,748 (19)	8,214 (18)
20-<40 MET hr/wk	35,268 (24)	7,637 (24)	21,588 (24)	11,040 (24)
40-<60 MET hr/wk	19,350 (13)	4,208 (13)	11,453 (13)	5,904 (13)
60+ MET hr/wk	37,549 (25)	8,261 (25)	19,318 (21)	10,264 (22)
Alcohol intake, n (%)				
never	7,739 (5)	1,751 (5)	7,870 (9)	4,445 (10)
Current occasionally to 1-3 times/month	22,216 (15)	5,064 (16)	24,243 (27)	12,584 (27)
1-2 times/week	38,272 (26)	8,304 (26)	22,884 (25)	11,639 (25)
3-4 times/week	40,456 (27)	8,757 (27)	19,180 (21)	9,435 (20)
daily or almost daily	38,826 (26)	8,551 (26)	16,284 (18)	8,014 (17)
Smoking status, n (%)				
Never	74,743 (51)	16,607 (51)	53,304 (59)	26,980 (59)
Former	55,302 (37)	12,048 (37)	30,094 (33)	15,971 (35)
Current	17,464 (12)	3,772 (12)	7,063 (8)	3,166 (7)

continued

Supplementary Table A-11 continued Distribution of baseline characteristics by missing status for any biomarker data for men and women

	Men		Women	
	No Missing for biomarker data	Any missing for biomarker data	No Missing for biomarker data	Any missing for biomarker data
	N=147,509	N=32,427	N=90,461	N=46,117
Total red and processed meat intake - times/week, median (25th - 75th percentiles)	4.0 (2.5-5.0)	4.0 (2.5-5.0)	2.5 (2.0-4.5)	2.5 (2.0-4.5)
Regular aspirin or ibuprofen use - yes, n (%)	38,959 (26)	8,585 (26)	21,040 (23)	10,997 (24)
Number of live births, median (25th - 75th percentiles)	-	-	2.0 (1.0-3.0)	2.0 (1.0-3.0)
Ever use of oral contraceptive pill - yes, n (%)	-	-	70,767 (78)	35,972 (78)
Ever use of hormone therapy - yes, n (%)	-	-	40,825 (45)	23,354 (51)
Adiposity measures Waist Circumference - cm, n (%)				
=94 M; <=80 F <	68,717 (47)	15,223 (47)	35,420 (39)	18,919 (41)
<=102 M; <=88 F	42,154 (29)	9,074 (28)	23,743 (26)	12,077 (26)
>102 M; >88 F	36,638 (25)	8,130 (25)	31,298 (35)	15,121 (33)
Waist Circumference - cm, median (25th - 75th percentiles)	95.0 (89.0-102.0)	95.0 (89.0-103.0)	84.0 (76.0-92.0)	83.0 (76.0-92.0)
Body Mass Index - kg/m2, n (%)				
<25	38,572 (26)	8,702 (27)	33,653 (37)	18,663 (40)
≥25-<30	74,924 (51)	16,256 (50)	35,827 (40)	17,794 (39)
≥30	34,013 (23)	7,469 (23)	20,981 (23)	9,660 (21)
Body Mass Index - kg/m2, median (25th - 75th percentiles)	27.1 (24.9-29.7)	27.1 (24.8-29.7)	26.3 (23.8-29.7)	26.0 (23.4-29.2)
Waist-Hip Ratio, n (%)				
<=0.89 M; <=0.76 F	40,028 (27)	8,828 (27)	20,621 (23)	10,438 (23)
<=0.98 M; <=0.86 F	74,935 (51)	16,375 (50)	46,810 (52)	23,686 (51)
>0.98 M; >0.86 F	32,546 (22)	7,224 (22)	23,030 (25)	11,993 (26)
Waist to Hip ratio, median (25th - 75th percentiles)	0.9 (0.9-1.0)	0.9 (0.9-1.0)	0.8 (0.8-0.9)	0.8 (0.8-0.9)

CSE Certificate of Secondary Education; GCSE General Certificate of Secondary Education; NVQ National Vocational Education; HND Higher National Diploma; HNC Higher National Certificate; MET Metabolic Equivalent of Task; W women; M men

Appendix B Supplementary material related to publication presented in Chapter 6

Supplementary Table B-1 Expressions used to estimate the interventional direct and indirect effect and their interpretations Notations: Y= outcome, ER-positive postmenopausal breast cancer; X=exposure, binary body mass index or waist circumference; Ma=mediator a (fasting insulin); Mb= mediator b (calculated free estradiol); C=baseline confounders not affected by exposure; Y_x, Ma_x, Mb_x = observed value of outcome or mediators when exposure X is set to value x (i.e. 0 when unexposed and 1 when exposed); Y_{xmamb} = observed value of outcome when exposure X is set to value x, Ma to ma and Mb to mb.

#	Effect	Expression [20] *
1	Interventional direct effect not through fasting insulin or calculated free estradiol	$\frac{E[\sum_{ma}\sum_{mb}E(Y_{1mamb}\mid c)P(Ma_0=ma,Mb_0=mb\mid c)]}{E[\sum_{ma}\sum_{mb}E(Y_{0mamb}\mid c)P(Ma_0=ma,Mb_0=mb\mid c)]}$
2	Interventional indirect effect through fasting insulin	$\frac{E[\sum_{ma}\sum_{mb}E(Y_{1mamb}\mid c)P(Ma_1=ma\mid c)\,P(Mb_0=mb\mid c)]}{E[\sum_{ma}\sum_{mb}E(Y_{1mamb}\mid c)P(Ma_0=ma\mid c)\,P(Mb_0=mb\mid c)]}$
3	Interventional indirect effect through calculated free estradiol	$\frac{E[\sum_{ma}\sum_{mb}E(Y_{1mamb}\mid c)P(Mb_1=mb\mid c)\,P(Ma_1=ma\mid c)]}{E[\sum_{ma}\sum_{mb}E(Y_{1mamb}\mid c)P(Mb_0=mb\mid c)\,P(Ma_1=ma\mid c)]}$
4	Interventional indirect effect through dependence of fasting insulin and calculated free estradiol	$\left(\frac{E[\sum_{ma}\sum_{mb}E(Y_{1mamb}\mid c)P(Ma_1=ma,Mb_1=mb\mid c)]}{E[\sum_{ma}\sum_{mb}E(Y_{0mamb}\mid c)P(Ma_0=ma,Mb_0=mb\mid c)]}\right)/$ $\left(\frac{E[\sum_{ma}\sum_{mb}E(Y_{1mamb}\mid c)P(Ma_0=ma,Mb_0=mb\mid c)]}{E[\sum_{ma}\sum_{mb}E(Y_{0mamb}\mid c)P(Ma_0=ma,Mb_0=mb\mid c)]}\right.$ $\times\frac{E[\sum_{ma}\sum_{mb}E(Y_{1mamb}\mid c)P(Ma_1=ma\mid c)\,P(Mb_1=mb\mid c)]}{E[\sum_{ma}\sum_{mb}E(Y_{1mamb}\mid c)P(Ma_0=ma\mid c)\,P(Mb_1=mb\mid c)]}$ $\left.\times\frac{E[\sum_{ma}\sum_{mb}E(Y_{1mamb}\mid c)P(Mb_1=mb\mid c)\,P(Ma_1=ma\mid c)]}{E[\sum_{ma}\sum_{mb}E(Y_{1mamb}\mid c)P(Mb_0=mb\mid c)\,P(Ma_1=ma\mid c)]}\right)$

*The corresponding interpretations for the effects are provided in Table 6-1.

Supplementary Table B-2 Results of multiple-mediator analysis with fasting insulin and free estradiol as the mediators of interest, body mass index (≥30 vs <25 kg/m^2), and non-estrogen receptor negative postmenopausal breast cancer as outcome.

		Risk ratio (95% confidence interval)		Risk difference, per 10,000 (95% confidence interval)	
	Effect	Complete case *	Multiple imputation†	Complete case *	Multiple imputation†
BMI ≥25 vs <25 kg/m^2	Total effect	1.14 (0.67 to 1.93)	1.27 (0.84 to 1.91)	15.4 (-39.9 to 70.7)	30.1 (-15.9 to 76.1)
	Interventional direct effect not through insulin and free estradiol	0.90 (0.46 to 1.74)	1.05 (0.64 to 1.74)	-9.6 (-69.1 to 50.0)	7.9 (-45.9 to 61.7)
	Interventional indirect effect through insulin	1.02 (0.84 to 1.23)	1.07 (0.89 to 1.29)	6.9 (-13.9 to 27.7)	13.8 (-9.9 to 37.4)
	Interventional indirect effect through free estradiol	1.28 (1.06 to 1.56)	1.18 (1.05 to 1.34)	27.8 (8.5 to 47.1)	23.0 (7.6 to 38.5)
	Interventional indirect effect through the interdependence of insulin and free estradiol	0.92 (0.80 to 1.05)	0.91 (0.79 to 1.04)	-8.4 (-25.8 to 9.1)	-12.3 (-33.6 to 9.1)
BMI Per 5 kg/m^2 ‡	Total effect	1.16 (0.90 to 1.49)	1.27 (1.05 to 1.53)	19.2 (-12.4 to 50.9)	34.4 (5.5 to 63.3)
	Interventional direct effect not through insulin and free estradiol	0.96 (0.70 to 1.31)	1.15 (0.89 to 1.47)	3.6 (-27.7 to 35.0)	20.7 (-15.4 to 56.9)
	Interventional indirect effect through insulin	0.98 (0.82 to 1.17)	1.00 (0.86 to 1.17)	1.8 (-23.3 to 26.9)	4.5 (-21.2 to 30.2)
	Interventional indirect effect through free estradiol	1.25 (1.07 to 1.45)	1.13 (1.01 to 1.27)	27.4 (7.4 to 47.3)	19.5 (2.6 to 36.5)
	Interventional indirect effect through the interdependence of insulin and free estradiol	0.97 (0.91 to 1.03)	0.96 (0.89 to 1.02)	-2.9 (-12.3 to 6.5)	-6.4 (-18.4 to 5.6)
Waist circumference >80 vs ≤80 cm	Total effect	1.10 (0.68 to 1.76)	1.47 (1.01 to 2.15)	11.2 (-40.3 to 62.7)	50.0 (4.6 to 94.9)
	Interventional direct effect not through insulin and free estradiol	0.72 (0.34 to 1.49)	1.12 (0.66 to 1.90)	-25.1 (-77.9 to 27.7)	15.6 (-36.1 to 67.2)
	Interventional indirect effect through insulin	1.10 (0.83 to 1.45)	1.15 (0.91 to 1.45)	16.2 (-12.4 to 44.8)	23.9 (-7.2 to 55.1)
	Interventional indirect effect through free estradiol	1.35 (1.09 to 1.69)	1.18 (1.02 to 1.36)	34.2 (10.8 to 57.7)	25.2 (5.1 to 45.4)
	Interventional indirect effect through the interdependence of insulin and free estradiol	0.95 (0.82 to 1.10)	0.93 (0.81 to 1.07)	-7.0 (-21.3 to 7.3)	-10.0 (-28.2 to 8.1)
Waist circumference Per 10 cm ‡	Total effect	1.11 (0.88 to 1.40)	1.30 (1.10 to 1.55)	12.9 (-15.2 to 40.9)	38.9 (13.2 to 64.6)
	Interventional direct effect not through insulin and free estradiol	0.94 (0.71 to 1.24)	1.17 (0.96 to 1.43)	0.3 (-24.4 to 25.0)	22.7 (-2.9 to 48.2)
	Interventional indirect effect through insulin	1.03 (0.89 to 1.19)	1.05 (0.93 to 1.18)	5.4 (-12.3 to 23.1)	8.9 (-9.1 to 26.9)
	Interventional indirect effect through free estradiol	1.18 (1.04 to 1.33)	1.08 (1.00 to 1.17)	19.2 (4.6 to 33.9)	12.1 (0.7 to 23.5)
	Interventional indirect effect through the interdependence of insulin and free estradiol	0.99 (0.94 to 1.04)	0.98 (0.93 to 1.02)	-1.7 (-7.0 to 3.6)	-3.1 (-9.7 to 3.5)

***** Analysis included 66 cases and 381 controls and models for mediation analyses were based on 600,000 Monte Carlo draws and 95% confidence intervals were constructed from 1,000 bootstrap draws. † Analysis included 111 cases 604 controls and mediation analyses were performed within each imputed dataset and resulting estimates and standard errors (derived from bootstrap draws) were pooled to obtain the multiple imputation estimates and 95% confidence intervals. All models were adjusted for country of birth, socioeconomic index for area, education, physical activity, smoking status, alcohol intake, history of cardiovascular disease, history of arthritis, age at menarche, history of hormonal contraceptive use, history of hormone therapy use, and parity and age at first live birth. ‡ These models assume that the total effect as well as the interventional direct and indirect effects are the same for every 5-unit increment in BMI or 10-unit increment in waist circumference.

Supplementary Table B-3 Baseline characteristics for all women eligible for mediation analysis versus women with complete data for fasting insulin and estradiol

	Women with missing data for fasting insulin or estradiol (n=438)	Women with complete data for fasting insulin and estradiol (n=740)	All women eligible for mediation analysis (n=1,178)
Breast cancer cases; n (%)	58 (13)	91 (12)	149 (13)
Age at baseline, *years*; median (interquartile range)	62 (57 – 66)	62 (57 – 65)	62 (57 – 65)
Country of birth, *grouped*; n (%)			
Australia/New Zealand/ Northern Europe	336 (77)	541 (73)	877 (74)
Southern Europe	102 (23)	199 (27)	301 (26)
Socioeconomic position, *quintile*; n (%)			
1 (most disadvantaged)	97 (22)	161 (22)	258 (22)
2	101 (23)	174 (24)	275 (23)
3	49 (11)	107 (15)	156 (13)
4	74 (17)	140 (19)	214 (18)
5 (least disadvantaged)	117 (27)	158 (21)	275 (23)
Physical activity score; n (%)			
None	82 (19)	159 (22)	241 (21)
Low	98 (22)	155 (21)	253 (22)
Medium	157 (36)	310 (42)	467 (40)
High	101 (23)	116 (16)	217 (18)
Smoking status; n (%)			
Never smoked	300 (69)	559 (76)	859 (73)
Former smoker	103 (24)	131 (18)	234 (20)
Current smoker	35 (8)	50 (7)	85 (7)
Alcohol consumed per day in the last decade, *grams per day*; n (%)			
0	223 (51)	408 (55)	631 (54)
1-19	171 (39)	277 (37)	448 (38)
³20	44 (10)	55 (7)	99 (8)
Personal history of cardiovascular disease; n (%)	23 (5)	47 (6)	70 (6)
Personal history of arthritis; n (%)	191 (44)	369 (50)	560 (48)
Age at menarche, *years*; n (%)			
<12	61 (14)	105 (14)	166 (14)
12	88 (20)	146 (20)	234 (20)
13	103 (24)	196 (27)	299 (25)
>13	186 (43)	293 (40)	479 (41)
Personal history of contraceptive pill use; n (%)	210 (48)	307 (42)	517 (44)
Personal history of hormone therapy use; n (%)	66 (15)	98 (13)	164 (14)
Parity and age at first live birth; n (%)			
Nulliparous	49 (11)	87 (12)	136 (12)
1	26 (6)	55 (7)	81 (7)
>1 & <25 years	190 (43)	291 (39)	481 (41)
>1 & ³25 years	173 (40)	307 (42)	480 (41)
Body mass index, *kg/m²*; median (interquartile range)	26.9 (24.1-30.5)	26.7 (24.2-30.1)	26.8 (24.2-30.3)
³18.5 - <25	145 (33)	247 (33)	392 (33)
³25 - <30	173 (40)	304 (41)	477 (41)
³30	120 (27)	189 (26)	309 (26)
Waist circumference, *cm*; median (interquartile range)	81 (74-90)	81 (74-90)	81 (74-90)
≤80	208 (48)	345 (47)	553 (47)
>80 - ≤88	113 (26)	184 (25)	297 (25)
>88	117 (27)	211 (29)	328 (28)

Appendix C Supplementary material related to publication presented in Chapter 7

Supplementary Table C-1 The models used for estimation of the effects, adiponectin, inflammatory markers, C-peptide, and free estradiol

	Effect	Natural indirect effect through all the biomarkers	Natural indirect effect through reduced adiponectin and increased inflammation	Natural indirect effect through reduced adiponectin levels	Natural indirect effect through increased inflammation excluding the potential influence of adiponectin	Natural indirect effect though increased C-peptide excluding the potential influences of adiponectin and inflammation	Natural indirect effect though estrogens excluding the potential influences of adiponectin, inflammation, and c-peptide	Natural direct effect not through any of the biomarkers
Mediator models; linear regression limited to the controls, for	**Adiponectin**, conditional on exposure and confounders	✓	✓	✓	✓	✓	✓	✓
	IL-6, conditional on adiponectin, exposure and confounders	✓	✓		✓	✓	✓	✓
	IL-1RA, conditional on adiponectin, IL-6, exposure and covariates	✓	✓		✓	✓	✓	✓
	TNF-R1, conditional on adiponectin, IL-6, IL-1RA, exposure and confounders	✓	✓		✓	✓	✓	✓
	TNF-R2, conditional on adiponectin, IL-6, IL-1RA, exposure and confounders	✓	✓		✓	✓	✓	✓
	CRP, conditional on adiponectin, IL-6, IL-1RA, TNF-R1, TNF-R2, exposure and confounders	✓	✓		✓	✓	✓	✓
	C-peptide, conditional on adiponectin, IL-6, IL-1RA, TNF-R1, TNF-R2, CRP, exposure and confounders	✓				✓	✓	✓
	Free estradiol, conditional on adiponectin, IL-6, IL-1RA, TNF-R1, TNF-R2, CRP, C-peptide, exposure and confounders	✓					✓	✓
	Estrone, conditional on free estradiol, adiponectin, IL-6, IL-1RA, TNF-R1, TNF-R2, CRP, C-peptide, exposure and confounders	✓					✓	✓
Outcome models; logistic regression conditional on exposure, confounders and	**Adiponectin**		✓	✓	✓	✓	✓	
	Adiponectin, **IL-6, IL-1RA, TNF-R1, TNF-R2, CRP**		✓		✓	✓	✓	
	Adiponectin, IL-6, IL-1RA, TNF-R1, TNF-R2, CRP, **C-peptide**					✓	✓	
	Adiponectin, IL-6, IL-1RA, TNF-R1, TNF-R2, CRP, C-peptide, **free estradiol, estrone**	✓						✓

Abbreviations: IL-6 interleukin 6; IL-1RA interleukin 1 receptor antagonist; TNF-R1 tumour necrosis factor-receptor 1; TNF-R2 tumour necrosis factor-receptor 2; CRP C-reactive protein Confounders included age at blood collection, country region, age at menarche, number of full-term pregnancies, history of oral contraceptive use, history of hormone therapy, physical activity, and smoking status.

Supplementary Table C-2 Baseline characteristics for women with and without missing data for biomarkers

	No missing biomarker data N=390	Missing biomarker data N=79
Endometrial Cancer Cases; n(%)	131 (34)	32 (41)
Age at blood collection, *years*; median (interquartile range)	60.1 (56.7-63.3)	59.4 (55.8-62.1)
Age at the end of follow up, *years*; median (IQR)	71.0 (65.6-76.6)	69.4 (62.1-76.1)
Fasting status at blood draw, *hours*; n(%)		
<3	207 (53)	32 (41)
3-6	67 (17)	16 (20)
>6	116 (30)	31 (39)
Region of recruitment; n(%)		
Western Europe	110 (28)	10 (13)
Northern Europe	148 (38)	37 (47)
Southern Europe	132 (34)	32 (41)
Full-term pregnancies, *number*; median (interquartile range)	2.0 (1.0-3.0)	2.0 (1.0-2.0)
Age at menarche, *years;* median (interquartile range)	13.0 (12.0-14.0)	13.0 (12.0-14.0)
Had history of oral contraceptive use; n(%)	132 (34)	24 (30)
Had history of hormone therapy; n(%)	65 (17)	18 (23)
Smoking status; n(%)		
never	260 (67)	48 (61)
former	73 (19)	14 (18)
current	57 (15)	17 (22)
Physical activity; n(%)		
Inactive to moderately inactive	139 (36)	31 (39)
moderately active to active	251 (64)	48 (61)
Body Mass Index, kg/m^2; n(%)		
<25	151 (39)	25 (32)
≥25 & <30	152 (39)	30 (38)
≥30	87 (22)	24 (30)
median (interquartile range)	26.3 (23.6-29.5)	27.2 (23.4-30.8)
Waist Circumference, *cm*; n(%)		
<=80	149 (38)	27 (34)
>80 & <=88	102 (26)	23 (29)
>88	139 (36)	29 (37)
median (interquartile range)	84.1 (77.0-92.8)	84.0 (79.0-96.0)
Waist-hip ratio; n(%)		
≤0.78	130 (33)	23 (29)
>0.78 & ≤0.84	139 (36)	27 (34)
>0.84	121 (31)	29 (37)
median (interquartile range)	0.8 (0.8-0.9)	0.8 (0.8-0.9)

Supplementary Table C-3 Exposure-mediator association; estimated ratios in geometric mean of biomarkers associated with body mass index and waist circumference; complete case analyses limited to controls (with complete biomarker data) and excluding women categorized as overweight

	Ratio of geometric means (95% confidence interval)								
	Body Mass Index ≥30 vs. <25 kg/m2			Waist circumference >88 vs. ≤80 cm			Waist-hip ratio >0.84 vs. ≤0.78 cm		
Biomarker	n=155			n=183			n=165		
Adiponectin	0.79	(0.68 to	0.91)	0.71	(0.61 to	0.81)	0.66	(0.57 to	0.76)
Interleukin-6	1.80	(1.44 to	2.26)	1.64	(1.37 to	1.97)	1.59	(1.29 to	1.96)
Interleukin-1 receptor antagonist	1.61	(1.10 to	2.36)	1.58	(1.18 to	2.12)	1.53	(1.10 to	2.14)
Tumour necrosis factor-α	1.11	(0.87 to	1.41)	1.04	(0.86 to	1.26)	1.09	(0.89 to	1.34)
Tumour necrosis factor-receptor 1	1.21	(1.13 to	1.30)	1.17	(1.09 to	1.25)	1.09	(1.01 to	1.18)
Tumour necrosis factor-receptor 2	1.18	(1.10 to	1.28)	1.12	(1.04 to	1.20)	1.08	(1.00 to	1.17)
C-reactive protein	2.65	(1.91 to	3.68)	2.36	(1.81 to	3.09)	2.22	(1.64 to	3.00)
C-peptide	1.51	(1.26 to	1.81)	1.58	(1.37 to	1.83)	1.54	(1.32 to	1.80)
Free estradiol	1.29	(1.11 to	1.49)	1.36	(1.20 to	1.53)	1.30	(1.15 to	1.47)
Estrone	1.15	(0.96 to	1.36)	1.14	(1.00 to	1.31)	1.10	(0.97 to	1.25)

Models were adjusted for age at blood collection, time at blood collection, fasting status at blood collection, recruitment centre, age at menarche, number of full-term pregnancies, history of oral contraceptive use, history of hormone therapy, physical activity, and smoking status.

Supplementary Table C-4 Estimated exposure-outcome and mediator-outcome associations; complete case analysis

Body Mass Index	**≥30 vs <25 (analysis excluded women categorized as overweight)**			
OR (95% CI)	2.75	(1.48 to	5.09)	
Waist circumference	**>88 vs ≤80 (analysis excluded women categorized as overweight)**			
OR (95% CI)	2.31	(1.35 to	3.95)	
Waist-hip ratio	**>0.84 vs ≤0.78 (analysis excluded women categorized as overweight)**			
OR (95% CI)	1.62	(0.90 to	2.94)	
Biomarker	**Per doubling concentration**			**P value for evidence against linearity**
Adiponectin				
Model 1 (adj for confounders + BMI)	0.68	(0.48 to	0.96)	0.17
Interleukin-6				
Model 1 (adj for confounders + BMI)	1.12	(0.85 to	1.47)	
Model 2 (adj for confounders + BMI + adiponectin)	1.06	(0.80 to	1.41)	0.58
Interleukin-1 receptor antagonist				
Model 1 (adj for confounders + BMI)	1.21	(1.04 to	1.41)	
Model 2 (adj for confounders + BMI + adiponectin)	1.19	(1.02 to	1.38)	0.06
Tumour necrosis factor-α				
Model 1 (adj for confounders + BMI)	0.99	(0.78 to	1.26)	
Model 2 (adj for confounders + BMI + adiponectin)	0.97	(0.76 to	1.25)	0.68
Tumour necrosis factor-receptor 1				
Model 1 (adj for confounders + BMI)	0.90	(0.41 to	1.94)	
Model 2 (adj for confounders + BMI + adiponectin)	0.87	(0.39 to	1.94)	0.58
Tumour necrosis factor-receptor 2				
Model 1 (adj for confounders + BMI)	0.82	(0.45 to	1.50)	
Model 2 (adj for confounders + BMI + adiponectin)	0.80	(0.44 to	1.48)	0.31
C-reactive protein				
Model 1 (adj for confounders + BMI)	1.06	(0.88 to	1.27)	
Model 2 (adj for confounders + BMI + adiponectin)	1.03	(0.86 to	1.24)	0.92
C-peptide				
Model 1 (adj for confounders + BMI)	1.22	(0.86 to	1.73)	
Model 2 (adj for confounders + BMI + adiponectin + IL1-RA)	1.06	(0.73 to	1.54)	0.54
Free estradiol				
Model 1 (adj for confounders + BMI)	1.48	(0.99 to	2.23)	
Model 2 (adj for confounders + BMI + adiponectin + IL1-RA + C-peptide)	1.34	(0.88 to	2.04)	0.18
Estrone				
Model 1 (adj for confounders + BMI)	1.98	(1.26 to	3.13)	
Model 2 (adj for confounders + BMI + adiponectin + IL1-RA + C-peptide)	1.95	(1.22 to	3.13)	0.71

Models were adjusted for age at blood collection, time at blood collection, fasting status at blood draw, recruitment centre, age at menarche, number of full-term pregnancies, history of oral contraceptive use, history of hormone therapy, physical activity, and smoking status. Models for biomarkers were additionally adjusted for body mass index continuous.

Supplementary Table C-5 Estimated natural direct and indirect effects using sequential mediation analysis, analyses excluded women categorized as overweight, complete case analysis

	Body Mass Index ≥30 vs. <25 kg/m2				Waist Circumference >88 vs. ≤80 cm			
Number of cases/ number of controls	83/155				105/183			
Effects	Odds Ratio (95% CI)			% mediated on log odds scale	Odds Ratio (95% CI)			% mediated on log odds scale
Total effect (estimated as the product of natural direct and indirect effects)	2.37	(1.07 to	5.23)		2.19	(1.20 to	4.01)	
Natural indirect effect through all the biomarkers	2.00	(0.99 to	4.04)	81%	1.63	(0.91 to	2.89)	62%
Natural indirect effect through reduced adiponectin and increased inflammation	1.52	(0.86 to	2.69)	49%	1.44	(0.88 to	2.36)	47%
Natural indirect effect through reduced adiponectin levels	1.22	(0.97 to	1.53)	23%	1.25	(0.97 to	1.61)	29%
Natural indirect effect through increased inflammation, beyond the potential influence of adiponectin	1.25	(0.74 to	2.10)	25%	1.15	(0.76 to	1.75)	18%
Natural indirect effect through increased c-peptide levels, beyond the potential influences of adiponectin and inflammation	1.07	(0.88 to	1.30)	8%	1.05	(0.89 to	1.24)	6%
Natural indirect effect through increased free estradiol and estrone levels, beyond the potential influences of adiponectin, inflammation, and c-peptide	1.23	(0.84 to	1.80)	24%	1.07	(0.84 to	1.37)	9%
Natural direct effect not through any of the biomarkers	1.18	(0.46 to	3.06)		1.35	(0.62 to	2.95)	

Confounders included age at blood collection, country region, age at menarche, number of full-term pregnancies, history of oral contraceptive use, history of hormone therapy, physical activity, and smoking status. The reported 95% CIs are constructed from 1,000 bootstrap draws.

Supplementary Table C-6 Estimated natural direct and indirect effects using sequential mediation analysis, continuous exposure variables

	Body Mass Index per 5 kg/m2				Waist Circumference per 10 cm			
	Complete Case		Multiple Imputation		Complete Case		Multiple Imputation	
Number of cases/ number of controls	131/259		163/306		131/259		163/306	
Effects	Odds Ratio (95% CI)	% mediated on log odds scale	Odds Ratio (95% CI)	% mediated on log odds scale	Odds Ratio (95% CI)	% mediated on log odds scale	Odds Ratio (95% CI)	% mediated on log odds scale
Total effect (estimated as the product of natural direct and indirect effects)	1.54 (1.17 to 2.03)		1.58 (1.24 to 2.03)		1.38 (1.09 to 1.73)		1.37 (1.13 to 1.67)	
Natural indirect effect through all the biomarkers	1.21 (0.95 to 1.53)	41%	1.21 (0.98 to 1.48)	41%	1.25 (1.02 to 1.54)	71%	1.25 (1.05 to 1.50)	71%
Natural indirect effect through reduced adiponectin and increased inflammation	1.10 (0.92 to 1.32)	27%	1.13 (0.96 to 1.34)	27%	1.09 (0.99 to 1.20)	57%	1.20 (1.04 to 1.39)	57%
Natural indirect effect through reduced adiponectin levels	1.06 (0.99 to 1.12)	15%	1.07 (1.01 to 1.14)	15%	1.09 (0.99 to 1.20)	33%	1.11 (1.02 to 1.21)	33%
Natural indirect effect through increased inflammation, beyond the potential influence of adiponectin	1.04 (0.88 to 1.24)	12%	1.06 (0.91 to 1.23)	12%	1.07 (0.94 to 1.21)	25%	1.08 (0.96 to 1.21)	25%
Natural indirect effect through increased c-peptide levels, beyond the potential influences of adiponectin and inflammation	1.01 (0.94 to 1.09)	0%	1.00 (0.94 to 1.06)	0%	1.02 (0.95 to 1.08)	2%	1.01 (0.96 to 1.06)	2%
Natural indirect effect through increased free estradiol and estrone levels, beyond the potential influences of adiponectin, inflammation, and c-peptide	1.08 (0.95 to 1.24)	14%	1.07 (0.95 to 1.19)	14%	1.06 (0.95 to 1.18)	12%	1.04 (0.96 to 1.13)	12%
Natural direct effect not through any of the biomarkers	1.28 (0.92 to 1.77)		1.31 (0.98 to 1.76)		1.10 (0.81 to 1.49)		1.09 (0.85 to 1.41)	

Confounders included age at blood collection, country region, age at menarche, number of full-term pregnancies, history of oral contraceptive use, history of hormone therapy, physical activity, and smoking status. For complete case analyses, the reported 95% CIs are constructed from 1,000 bootstrap draws. For multiple imputation analysis mediation analyses were performed within each of the 20 imputed datasets and resulting estimates and standard errors (derived from 1,000 bootstrap draws) were pooled to obtain the multiple imputation estimates and 95% confidence intervals.

These models assume that the NDE and NIEs are the same for every 5-unit increment in BMI or 10-unit increment in waist circumference.

Supplementary Table C-7 Estimated natural direct and indirect effects using sequential mediation analysis, excluding women diagnosed with endometrial cancer within two years after recruitment

	Body Mass Index ≥30 vs. <25 kg/m2							Waist Circumference >88 vs. ≤80 cm						
	Complete Case			Multiple Imputation				Complete Case			Multiple Imputation			
Number of cases/ number of controls	59/155			69/185				75/183			88/212			
Effects	Odds Ratio (95% CI)			Odds Ratio (95% CI)			% mediated on log odds scale	Odds Ratio (95% CI)			Odds Ratio (95% CI)			% mediated on log odds scale
Total effect (estimated as the product of natural direct and indirect effects)	2.45	(0.93 to	6.44)	2.54	(1.11 to	5.82)		2.18	(1.02 to	4.67)	2.16	(1.13 to	4.13)	
Natural indirect effect through all the biomarkers	2.24	(0.93 to	5.43)	2.15	(0.97 to	4.78)	82%	1.71	(0.82 to	3.56)	1.90	(0.98 to	3.66)	83%
Natural indirect effect through reduced adiponectin and increased inflammation	1.43	(0.76 to	2.71)	1.50	(0.79 to	2.84)	43%	1.40	(0.78 to	2.51)	1.63	(0.94 to	2.80)	63%
Natural indirect effect through reduced adiponectin levels	1.27	(0.96 to	1.68)	1.42	(1.06 to	1.91)	38%	1.26	(0.95 to	1.67)	1.36	(1.02 to	1.81)	40%
Natural indirect effect through increased inflammation, beyond the potential influence of adiponectin	1.13	(0.64 to	2.00)	1.05	(0.61 to	1.82)	6%	1.11	(0.68 to	1.81)	1.20	(0.77 to	1.86)	24%
Natural indirect effect through increased c-peptide levels, beyond the potential influences of adiponectin and inflammation	1.10	(0.89 to	1.35)	1.07	(0.89 to	1.30)	8%	1.06	(0.88 to	1.28)	1.04	(0.88 to	1.23)	5%
Natural indirect effect through increased free estradiol and estrone levels, beyond the potential influences of adiponectin, inflammation, and c-peptide	1.43	(0.79 to	2.57)	1.34	(0.86 to	2.10)	31%	1.16	(0.85 to	1.58)	1.12	(0.83 to	1.52)	15%
Natural direct effect not through any of the biomarkers	1.09	(0.34 to	3.49)	1.18	(0.42 to	3.35)		1.27	(0.46 to	3.52)	1.14	(0.47 to	2.75)	

Confounders included age at blood collection, country region, age at menarche, number of full-term pregnancies, history of oral contraceptive use, history of hormone therapy, physical activity, and smoking status. For complete case analyses, the reported 95% CIs are constructed from 1,000 bootstrap draws. For multiple imputation analysis mediation analyses were performed within each of the 20 imputed datasets and resulting estimates and standard errors (derived from 1,000 bootstrap draws) were pooled to obtain the multiple imputation estimates and 95% confidence intervals.

Appendix D Supplementary material related to publication presented in Chapter 8

Supplementary Table D-1 Pearson correlation coefficients between potential confounders not included in the final models due to missing values and exospore and mediator variables

	Adiposity measures			Biomarkers (mediators)								
Potential confounders	BMI	Waist circumference	Waist-hip ratio	Adiponectin	Leptin	Resistin	HGF	PAI	TNF-α	IL-6	CRP	Fasting insulin	Estradiol
Age at menopause	-0.17	-0.16	-0.03	0.25	-0.12	-0.09	0.02	-0.24	0.04	0.05	0.02	-0.23	-0.09
Red meat intake	0.10	0.06	-0.08	0.04	0.03	-0.01	0.07	0.25	-0.10	0.18	0.28	-0.01	0.01
Family history breast cancer	0.02	0.06	0.10	0.04	0.06	-0.05	-0.02	0.10	0.07	0.07	0.04	0.08	-0.14
Family history endometrial cancer	-0.03	-0.05	-0.12	-0.04	-0.13	-0.06	0.03	0.12	0.15	0.03	0.17	0.05	-0.08
Family history colorectal cancer	-0.18	-0.08	0.11	-0.02	-0.09	0.02	0.20	0.06	0.01	0.07	0.13	-0.12	0.00

Abbreviations BMI Body Mass Index; HGF hepatocyte growth factor; PAI plasminogen activator inhibitor-1; TNF-α tumour necrosis factor-α; IL-6 interleukin-6; CRP C-reactive protein

Supplementary Table D-2 The models used for estimation of the interventional indirect and direct effects

		Effect				
	Model	Interventional indirect effect through all the biomarkers	Interventional indirect effect through inflammatory status	Interventional indirect effect through fasting insulin	Interventional indirect effect through estradiol	Interventional direct effect
Mediator models; linear regression limited to the participants in the sub-cohort	**Adiponectin**, conditional on exposure and confounders	✓	✓	✓	✓	✓
	Leptin, conditional on adiponectin, exposure and confounders	✓	✓	✓	✓	✓
	Resistin, conditional on adiponectin, leptin, exposure and confounders	✓	✓	✓	✓	✓
	HGF, conditional on adiponectin, leptin, resistin, exposure and confounders	✓	✓	✓	✓	✓
	PAI-1, conditional on adiponectin, leptin, resistin, HGF, exposure and confounders	✓	✓	✓	✓	✓
	TNF-α, conditional on adiponectin, leptin, resistin, HGF, PAI-1, exposure and confounders	✓	✓	✓	✓	✓
	IL-6, conditional on adiponectin, leptin, resistin, HGF, PAI-1, TNF-α, exposure and confounders	✓	✓	✓	✓	✓
	CRP, conditional on adiponectin, leptin, resistin, HGF, PAI-1, TNF-α, IL-6 exposure and confounders	✓	✓	✓	✓	✓
	Fasting insulin, conditional on exposure and confounders		✓	✓	✓	
	Fasting insulin, conditional on adiponectin, leptin, resistin, HGF,PAI-1, TNF-α, IL-6, CRP, exposure and confounders	✓				✓
	Estradiol, conditional on exposure and confounders		✓	✓	✓	
	Esttradiol, conditional on adiponectin, leptin, resistin, HGF,PAI-1, TNF-α, IL-6, CRP, fasting insulin, exposure and confounders	✓				✓
	Logistic regression model for the outcome conditional on adiponectin, leptin, resistin, HGF,PAI-1, TNF-α, IL-6, CRP, fasting insulin, estradiol, exposure, and confounders	✓	✓	✓	✓	✓

Abbreviations: HGF Hepatocyte growth factor; PAI-1 Plasminogen activator inhibitor-1; TNF-α Tumour necrosis factor-α; IL-6 Interleukin-6; CRP C-reactive protein

Supplementary Table D-3 Interventional indirect and direct effects

	Effects	Waist circumference >88 vs ≤80 cm Risk Ratio (95% CI)	Waist circumference >88 vs ≤80 cm Risk Difference per 10,000 (95% CI)	BMI ≥30 vs ≥18.5-<25 kg/m2 Risk Ratio (95% CI)	BMI ≥30 vs ≥18.5-<25 kg/m2 Risk Difference per 10,000 (95% CI)	Waist-hip ratio >0.83 vs ≤0.77 Risk Ratio (95% CI)	Waist-hip ratio >0.83 vs ≤0.77 Risk Difference per 10,000 (95% CI)
ER-positive Breast Cancer	Total causal effect*	1.68 (0.96 to 2.79)	29.99 (1.35 to 62.19)	2.35 (1.22 to 4.34)	49.88 (15.73 to 101.92)	1.82 (0.93 to 3.35)	32.66 (2.68 to 69.85)
	Interventional indirect effect through all biomarkers	2.09 (1.15 to 3.46)	38.66 (16.45 to 67.28)	3.82 (1.71 to 7.69)	64.15 (38.51 to 114.24)	1.55 (0.93 to 2.70)	25.62 (1.15 to 55.57)
	Interventional indirect effect through Adiponectin	0.90 (0.73 to 1.06)	-3.66 (-14.42 to 1.07)	0.84 (0.62 to 1.01)	-3.73 (-16.13 to -0.44)	0.95 (0.73 to 1.21)	-2.21 (-18.88 to 7.43)
	Interventional indirect effect through Leptin	1.11 (0.69 to 1.76)	3.41 (-12.19 to 20.80)	1.30 (0.61 to 2.50)	5.71 (-8.17 to 25.60)	0.92 (0.68 to 1.22)	-3.53 (-22.75 to 7.50)
	Interventional indirect effect through Resistin	0.97 (0.83 to 1.03)	-1.02 (-7.37 to 0.64)	0.94 (0.75 to 1.01)	-1.58 (-8.65 to 0.06)	1.00 (0.96 to 1.06)	0.01 (-1.89 to 2.59)
	Interventional indirect effect through HGF	1.01 (0.96 to 1.12)	0.35 (-1.26 to 4.26)	1.06 (0.98 to 1.29)	1.41 (-0.25 to 9.19)	1.01 (0.96 to 1.14)	0.43 (-1.55 to 7.00)
	Interventional indirect effect through PAI	1.20 (0.91 to 1.57)	6.99 (-1.04 to 20.90)	1.22 (0.88 to 1.73)	5.55 (-1.13 to 19.96)	1.16 (0.95 to 1.51)	6.91 (-1.55 to 24.75)
	Interventional indirect effect through Il-6, CRP	0.93 (0.69 to 1.29)	-2.79 (-19.37 to 8.48)	1.28 (0.87 to 1.99)	8.48 (-2.04 to 29.77)	0.87 (0.64 to 1.17)	-6.61 (-27.71 to 5.11)
	Interventional indirect effect through fasting insulin	1.67 (1.08 to 2.65)	25.83 (9.71 to 50.85)	1.87 (1.12 to 3.26)	33.78 (11.22 to 77.89)	1.66 (1.16 to 2.84)	28.88 (10.21 to 71.31)
	Interventional indirect effect through estradiol	1.14 (1.02 to 1.39)	8.88 (1.80 to 24.66)	1.20 (1.02 to 1.60)	14.56 (2.87 to 42.71)	1.04 (0.97 to 1.24)	2.77 (-2.31 to 18.50)
	Interventional direct effect	0.80 (0.39 to 1.71)	-8.68 (-32.94 to 21.76)	0.62 (0.24 to 1.61)	-14.28 (-36.37 to 10.33)	1.18 (0.52 to 2.50)	7.03 (-21.27 to 44.27)
	Interdependence	1.00 (1.00 to 1.04)	0.69 (-1.97 to 7.41)	1.00 (1.00 to 1.04)	-0.01 (-9.48 to 12.90)	1.00 (0.95 to 1.01)	-1.03 (-9.87 to 3.35)

continued

Supplementary Table D-3 continued Interventional indirect and direct effects

	Effects	Waist circumference >88 vs ≤80 cm Risk Ratio (95% CI)	Waist circumference >88 vs ≤80 cm Risk Difference per 10,000 (95% CI)	BMI ≥30 vs ≥18.5-<25 kg/m2 Risk Ratio (95% CI)	BMI ≥30 vs ≥18.5-<25 kg/m2 Risk Difference per 10,000 (95% CI)	Waist-hip ratio >0.83 vs ≤0.77 Risk Ratio (95% CI)	Waist-hip ratio >0.83 vs ≤0.77 Risk Difference per 10,000 (95% CI)
Endometrial cancer	Total causal effect*	2.24 (1.03 to 5.03)	36.67 (9.17 to 82.91)	2.38 (0.97 to 5.44)	48.25 (10.56 to 122.02)		
	Interventional indirect effect through all biomarkers	2.01 (0.86 to 3.97)	33.36 (6.61 to 67.52)	3.16 (1.04 to 8.26)	56.99 (23.81 to 116.94)		
	Interventional indirect effect through Adiponectin	1.01 (0.79 to 1.35)	0.30 (-8.03 to 11.14)	1.02 (0.82 to 1.38)	0.50 (-4.20 to 11.76)		
	Interventional indirect effect through Leptin	1.38 (0.73 to 2.79)	12.53 (-7.20 to 67.88)	2.09 (0.80 to 5.22)	27.79 (0.32 to 144.59)		
	Interventional indirect effect through Resistin	1.07 (0.96 to 1.34)	3.26 (-0.65 to 20.72)	1.00 (0.79 to 1.18)	-0.09 (-13.43 to 11.79)		
	Interventional indirect effect through HGF	1.03 (0.96 to 1.25)	1.53 (-1.62 to 18.05)	1.07 (0.90 to 1.44)	3.99 (-3.96 to 46.19)		
	Interventional indirect effect through PAI	1.02 (0.73 to 1.47)	1.17 (-19.19 to 27.14)	0.97 (0.58 to 1.58)	-1.61 (-59.75 to 27.49)		
	Interventional indirect effect through Il-6, CRP	0.89 (0.52 to 1.43)	-5.79 (-59.51 to 12.09)	0.90 (0.44 to 1.50)	-5.81 (-85.98 to 16.35)		
	Interventional indirect effect through fasting insulin	1.13 (0.61 to 1.83)	5.89 (-28.71 to 25.91)	1.22 (0.60 to 2.32)	10.75 (-35.98 to 43.76)		
	Interventional indirect effect through estradiol	1.27 (1.04 to 1.69)	14.05 (3.81 to 37.67)	1.30 (0.99 to 2.05)	18.41 (0.43 to 57.06)		
	Interventional direct effect	1.11 (0.43 to 3.75)	3.30 (-19.47 to 55.14)	0.75 (0.21 to 3.11)	-8.74 (-36.57 to 42.08)		
	Interdependence	1.00 (1.00 to 1.02)	0.42 (-2.76 to 10.91)	1.00 (1.00 to 1.01)	3.06 (-5.34 to 28.67)		

continued

Supplementary Table D-3 continued Interventional indirect and direct effects

	Effects	Waist circumference >88 vs ≤80 cm Risk Ratio (95% CI)	Waist circumference >88 vs ≤80 cm Risk Difference per 10,000 (95% CI)	BMI ≥30 vs ≥18.5-<25 kg/m2 Risk Ratio (95% CI)	BMI ≥30 vs ≥18.5-<25 kg/m2 Risk Difference per 10,000 (95% CI)	Waist-hip ratio >0.83 vs ≤0.77 Risk Ratio (95% CI)	Waist-hip ratio >0.83 vs ≤0.77 Risk Difference per 10,000 (95% CI)
Colorectal cancer	Total causal effect*	1.61 (0.94 to 2.58)	29.23 (1.46 to 64.34)	1.76 (0.95 to 3.06)	35.57 (2.29 to 77.44)		
	Interventional indirect effect through all biomarkers	1.61 (0.95 to 2.52)	29.06 (2.12 to 56.70)	2.02 (1.06 to 4.00)	41.64 (12.76 to 80.24)		
	Interventional indirect effect through Adiponectin	1.01 (0.86 to 1.19)	0.69 (-7.60 to 9.82)	0.94 (0.75 to 1.10)	-2.46 (-15.18 to 2.97)		
	Interventional indirect effect through Leptin	1.22 (0.78 to 1.92)	10.52 (-11.32 to 42.23)	1.39 (0.73 to 2.56)	14.66 (-10.99 to 53.74)		
	Interventional indirect effect through Resistin	1.06 (1.00 to 1.21)	3.84 (0.07 to 15.55)	1.04 (0.97 to 1.20)	2.29 (-0.91 to 12.29)		
	Interventional indirect effect through HGF	0.99 (0.91 to 1.05)	-0.66 (-7.96 to 2.32)	0.99 (0.87 to 1.06)	-0.75 (-9.47 to 2.72)		
	Interventional indirect effect through PAI	1.22 (0.97 to 1.61)	13.70 (-0.99 to 43.78)	1.16 (0.86 to 1.57)	8.77 (-5.75 to 37.32)		
	Interventional indirect effect through Il-6, CRP	0.86 (0.65 to 1.15)	-10.52 (-41.78 to 6.28)	0.87 (0.59 to 1.23)	-8.08 (-43.15 to 8.80)		
	Interventional indirect effect through fasting insulin	1.16 (0.77 to 1.69)	10.62 (-22.40 to 34.00)	1.50 (0.88 to 2.39)	27.44 (-3.05 to 63.16)		
	Interventional indirect effect through estradiol	1.01 (0.89 to 1.14)	0.92 (-9.97 to 9.42)	0.99 (0.83 to 1.17)	-0.60 (-17.73 to 12.16)		
	Interventional direct effect	1.00 (0.51 to 1.94)	0.18 (-26.36 to 35.82)	0.87 (0.39 to 2.03)	-6.07 (-33.73 to 32.90)		
	Interdependence	1.00 (0.97 to 1.00)	-0.06 (-1.30 to 3.44)	1.00 (0.98 to 1.00)	0.38 (-1.11 to 7.05)		

Abbreviations HGF Hepatocyte growth factor; PAI-1 Plasminogen activator inhibitor-1; TNF-α Tumour necrosis factor-α; IL-6 Interleukin-6; CRP C-reactive protein
All analyses were complete case. Models were adjusted for age at baseline, educational attainment, alcohol intake, ever smoked, physical activity, NSAID use, age at menarche, parity, oral contraceptive use, hormone therapy use.*Total effects presented here were estimated as the product of interventional indirect and direct effects.